OM protein - protein search, using sw model

Run on: June 2, 2003, 08:39:40 ; Search time 24 Seconds
(without alignments)
218.220 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	480	54.2	394	4	US-09-134-001C-4834	Sequence 4834, Ap
2	472	53.3	335	4	US-08-987-146-2	Sequence 2, Appli
3	472	53.3	419	4	US-09-120-426-2	Sequence 2, Appli
4	463	52.3	437	2	US-08-883-515-4	Sequence 4, Appli
5	429	48.5	433	2	US-08-883-515-2	Sequence 2, Appli
6	271	30.6	223	4	US-09-120-426-4	Sequence 4, Appli
7	196	22.1	283	4	US-08-961-083-124	Sequence 124, App
8	90.5	10.2	2285	4	US-09-308-375-2	Sequence 2, Appli
9	79	8.9	388	4	US-09-134-001C-4951	Sequence 4951, Ap
10	77	8.7	271	4	US-09-189-527-11	Sequence 11, Appl
11	76.5	8.6	422	4	US-09-134-001C-3034	Sequence 3034, Ap
12	74	8.4	327	4	US-09-134-001C-3477	Sequence 3477, Ap
13	74	8.4	1529	4	US-09-134-001C-3945	Sequence 3945, Ap
14	73.5	8.3	450	2	US-08-861-464-2	Sequence 2, Appli
15	73.5	8.3	450	2	US-08-396-001-2	Sequence 2, Appli
16	73.5	8.3	450	4	US-09-323-433A-2	Sequence 2, Appli
17	73	8.2	1026	1	US-08-194-290-7	Sequence 7, Appli
18	73	8.2	1026	2	US-08-614-377A-7	Sequence 7, Appli
19	73	8.2	1026	4	US-09-142-648B-7	Sequence 7, Appli
20	71.5	8.1	267	2	US-08-557-128-4	Sequence 4, Appli
21	71.5	8.1	267	4	US-09-242-690A-36	Sequence 36, Appl
22	71.5	8.1	1004	4	US-09-268-347-30	Sequence 30, Appl
23	71	8.0	267	4	US-09-302-620B-106	Sequence 106, App
24	71	8.0	267	6	5204252-2	Patent No. 5204252
25	71	8.0	445	2	US-08-935-450-4	Sequence 4, Appli
26	71	8.0	1451	1	US-08-308-872B-4	Sequence 4, Appli
27	70.5	8.0	348	4	US-09-134-001C-5513	Sequence 5513, Ap
28	70	7.9	396	4	US-08-861-774E-84	Sequence 84, Appl
29	70	7.9	2647	2	US-08-583-562B-8	Sequence 8, Appli
30	70	7.9	2647	2	US-08-779-113-8	Sequence 8, Appli
31	69.5	7.9	355	5	PCT-US95-03866-32	Sequence 32, Appl
32	69	7.8	345	4	US-09-134-001C-3774	Sequence 3774, Ap
33	68.5	7.7	270	4	US-09-013-881-6	Sequence 6, Appli
34	68.5	7.7	459	4	US-09-491-785-2	Sequence 2, Appli
35	68	7.7	2314	4	US-09-268-347-49	Sequence 49, Appl
36	67.5	7.6	409	2	US-08-743-130A-2	Sequence 2, Appli
37	67.5	7.6	3816	4	US-09-428-517-3	Sequence 3, Appli
38	67	7.6	458	4	US-08-618-485B-1	Sequence 1, Appli
39	67	7.6	458	6	5177002-1	Patent No. 5177002
40	67	7.6	458	6	5177002-2	Patent No. 5177002
41	67	7.6	474	1	US-08-222-619-5	Sequence 5, Appli
42	67	7.6	474	5	PCT-US95-04075-5	Sequence 5, Appli
43	67	7.6	568	1	US-07-732-242C-3	Sequence 3, Appli
44	67	7.6	641	4	US-08-836-567-10	Sequence 10, Appl
45	67	7.6	3739	3	US-09-320-878-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-134-001C-4834
; Sequence 4834, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4834
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4834

Query Match 54.2%; Score 480; DB 4; Length 394;
Best Local Similarity 54.3%; Pred. No. 2.8e-47;
Matches 95; Conservative 33; Mismatches 47; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              :  ||||  ||| |||  | |  || :|: :|:|||||: ||:||::||: :  |:  :|
Db         35 MNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQ 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| |||||| |||||||| :|:|:| ||||||: ||  ||  |   | |||: :
Db         95 GADMVFVTAGMGGGTGTGAAPVVAKIAKEMGALTVGVVTRPFGFEGRKRQTQAAAGVESM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                  |||||:||  || : |||| |:||: ||:|| :||:|::|||   | :||
Db        155 KAAVDTLIVIPNDRLLDIVDKSTPMMEAFKEADNVLRQGVQGISDLIAVSGEVNL 209
```

RESULT 2
US-08-987-146-2
; Sequence 2, Application US/08987146
; Patent No. 6350866
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: Peery, Robert B.
; APPLICANT: Rockey, Pamela K.
; APPLICANT: Wang, Q. May
; APPLICANT: Rosteck Jr., Paul R.
; TITLE OF INVENTION: Streptococcus Pneumoniae Gene Sequence

```
;    TITLE OF INVENTION:  FtsZ
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/987,146
;      FILING DATE:  December 8, 1997
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11755
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  335 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-987-146-2

  Query Match             53.3%;  Score 472;  DB 4;  Length 335;
  Best Local Similarity   52.6%;  Pred. No. 1.8e-46;
  Matches   91;  Conservative   35;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          6 GVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDS 65
              ||||| |||| ||| : |  | ||  :|:||||| :||:|:::||: :  :  :  :
Db         38 GVEFIAANTDVQALSSTKAETVIQLGPKLTRGLGAGGQPEVGRKAAEESEETLTEAISGA 97

Qy         66 NMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAK 125
              :|:||| ||||| |||||| :| :|::|| |||||| || || | : |  |: :| :
Db         98 DMVFITAGMGGGSGTGAAPVIARIAKDLGALTVGVVTRPFGFEGSKRGQFAVEGINQLRE 157

Qy        126 YVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :||||::: | ||| : || | :|||  ||:|| :||:|:|||| ||||||
Db        158 HVDTLLIISNNNLLEIVDKKTPLLEALSEADNVLRQGVQGITDLITNPGLINL 210


RESULT 3
US-09-120-426-2
; Sequence 2, Application US/09120426
; Patent No. 6197300
; GENERAL INFORMATION:
;  APPLICANT: Fueyo, Joanna Lynn
;  APPLICANT: Lonetto, Michael A.
;  APPLICANT: Pearce, Kenneth
;  TITLE OF INVENTION: ftsZ
;  FILE REFERENCE: GM10068
;  CURRENT APPLICATION NUMBER: US/09/120,426
;  CURRENT FILING DATE: 1998-07-22
;  EARLIER APPLICATION NUMBER: 60/055,720
;  EARLIER FILING DATE: 1997-08-12
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Streptococcus pneumoniae
US-09-120-426-2

  Query Match             53.3%;  Score 472;  DB 4;  Length 419;
  Best Local Similarity   52.6%;  Pred. No. 2.6e-46;
  Matches   91;  Conservative   35;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          6 GVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDS 65
              ||||| |||| ||| : |  | ||  :|:||||| :||:|:::||: :  :  :  :
Db         38 GVEFIAANTDVQALSSTKAETVIQLGPKLTRGLGAGGQPEVGRKAAEESEETLTEAISGA 97

Qy         66 NMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAK 125
              :|:||| ||||| |||||| :| :|::|| |||||| || || | : |  |: :| :
Db         98 DMVFITAGMGGGSGTGAAPVIARIAKDLGALTVGVVTRPFGFEGSKRGQFAVEGINQLRE 157

Qy        126 YVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :||||::: | ||| : || | :|||  ||:|| :||:|:|||| ||||||
Db        158 HVDTLLIISNNNLLEIVDKKTPLLEALSEADNVLRQGVQGITDLITNPGLINL 210


RESULT 4
US-08-883-515-4
; Sequence 4, Application US/08883515
; Patent No. 5981836
;  GENERAL INFORMATION:
;    APPLICANT:  Osteryoung, Katherine W
;    TITLE OF INVENTION:  PLANT CHLOROPLAST DIVISION GENES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  1 South Pinckney Street
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  US
;      ZIP:  53701-2113
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,515
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seay, Nicholas J.
;      REGISTRATION NUMBER:  27,386
;      REFERENCE/DOCKET NUMBER:  920905.90016
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 437 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-883-515-4

  Query Match             52.3%;  Score 463;  DB 2;  Length 437;
  Best Local Similarity   49.7%;  Pred. No. 3e-45;
  Matches   89;  Conservative   36;  Mismatches   52;  Indels    2;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRS--LAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQ 59
              |:: |||| : ||| ||:  |  |  ::: :||:|:|||||  ||:|  :| : |  |:
Db         99 SEMSGVEFWIVNTDIQAMRMSPVLPDNRLQIGKELTRGLGAGGNPEIGMNAARESKEVIE 158

Qy         60 RMLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAG 119
                |  |:|:|:| |||||| ||||||:| :|: :|||||: :|||  ||  ||  |  |
Db        159 EALYGSDMVFVTAGMGGGTGTGAAPVIAGIAKAMGILTVGIATTPFSFEGRRRTVQAQEG 218

Qy        120 VKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :  |   ||||||:||  ||    :|| : |||  |||:| :||:|::|:|  |||:|:
Db        219 LASLRDNVDTLIVIPNDKLLTAVSQSTPVTEAFNLADDILRQGVRGISDIITIPGLVNV 277


RESULT 5
US-08-883-515-2
```

```
; Sequence 2, Application US/08883515
; Patent No. 5981836
;  GENERAL INFORMATION:
;    APPLICANT:  Osteryoung, Katherine W
;    TITLE OF INVENTION:  PLANT CHLOROPLAST DIVISION GENES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  1 South Pinckney Street
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  US
;      ZIP:  53701-2113
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,515
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seay, Nicholas J.
;      REGISTRATION NUMBER:  27,386
;      REFERENCE/DOCKET NUMBER:  920905.90016
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 433 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-883-515-2

  Query Match             48.5%;  Score 429;  DB 2;  Length 433;
  Best Local Similarity   48.3%;  Pred. No. 2.6e-41;
  Matches   86;  Conservative   40;  Mismatches   52;  Indels    0;  Gaps    0;

Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              :| |: |:|   ||| ||| :  | : : :|: :|:||| |  | ||:::||: |  |
Db         94 SSGLQSVDFYAINTDSQALLQFSAENPLQIGELLTRGLGTGGNPLLGEQAAEESKDAIAN 153

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
               |: |:::||| |||||| :||||||| :::: | |||||: ||  ||  |:  |   :
Db        154 ALKGSDLVFITAGMGGGTGSGAAPVVAQISKDAGYLTVGVVTYPFSFEGRKRSLQALEAI 213

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              ::| | ||||||:||  || :||: | : :||  ||||| :||:|::|:|  |||:|:
Db        214 EKLQKNVDTLIVIPNDRLLDIADEQTPLQDAFLLADDVLRQGVQGISDIITIPGLVNV 271


RESULT 6
US-09-120-426-4
; Sequence 4, Application US/09120426
; Patent No. 6197300
; GENERAL INFORMATION:
;  APPLICANT: Fueyo, Joanna Lynn
;  APPLICANT: Lonetto, Michael A.
;  APPLICANT: Pearce, Kenneth
;  TITLE OF INVENTION: ftsz
;  FILE REFERENCE: GM10068
;  CURRENT APPLICATION NUMBER: US/09/120,426
;  CURRENT FILING DATE: 1998-07-22
;  EARLIER APPLICATION NUMBER: 60/055,720
;  EARLIER FILING DATE: 1997-08-12
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Streptococcus pneumoniae
US-09-120-426-4

  Query Match             30.6%;  Score 271;  DB 4;  Length 223;
  Best Local Similarity   53.6%;  Pred. No. 1.9e-23;
  Matches   52;  Conservative   20;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         82 AAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAKYVDTLIVVPNQNLLAL 141
              :|||:| :|::|| |||||: ||  ||  | : |  |: :| ::||||::: | ||| :
Db          1 SAPVIARIAKDLGALTVGVVTRPFGFEGSKRGQFAVEGIDQLREHVDTLLIISNNNLLEI 60

Qy        142 ADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                || | :|||   ||:|| :||:|:|||| ||||||
Db         61 VDKKTPLLEALSEADNVLRQGVQGITDLITNPGLINL 97


RESULT 7
US-08-961-083-124
; Sequence 124, Application US/08961083
; Patent No. 6159469
;  GENERAL INFORMATION:
;    APPLICANT:  Choi et. al.
;    TITLE OF INVENTION:  Streptococcus pneumoniae Antigens and Vaccines
;    NUMBER OF SEQUENCES:  452
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,083
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  124:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  283 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-961-083-124

  Query Match             22.1%;  Score 196;  DB 4;  Length 283;
  Best Local Similarity   51.4%;  Pred. No. 1.3e-14;
  Matches   38;  Conservative   14;  Mismatches   22;  Indels    0;  Gaps    0;

Qy        105 FRSEGPNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVK 164
              |  ||  | : |  |: :| ::||||::: | ||| : || | :|||   ||:|| :||:
Db          1 FGFEGSKRGQFAVEGINQLREHVDTLLIISNNNLLEIVDKKTPLLEALSEADNVLRQGVQ 60

Qy        165 GVTDLIVRPGLINL 178
              |:|||  ||||||
Db         61 GITDLITNPGLINL 74


RESULT 8
```

```
US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
;  APPLICANT: Genencor International, Inc.
;  TITLE OF INVENTION: Proteases From Gram-Positive Organisms
;  FILE REFERENCE: GC394-PCT
;  CURRENT APPLICATION NUMBER: US/09/308,375
;  CURRENT FILING DATE: 1999-05-14
;  EARLIER APPLICATION NUMBER: EP9719636.4
;  EARLIER FILING DATE: 1997-09-15
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 2285
;   TYPE: PRT
;   ORGANISM: Bacillus subtilis
US-09-308-375-2

  Query Match             10.2%;  Score 90.5;  DB 4;  Length 2285;
  Best Local Similarity   21.4%;  Pred. No. 0.47;
  Matches   42;  Conservative   32;  Mismatches   79;  Indels   43;  Gaps    7;

Qy          1 ASQL-EGVEFIVANTDCQALGRSLAPHKI-----------------TLGKDITK------ 36
              | || |  :|| ||: :::|:|:: : |                 | || ::
Db       2091 AKQLNEFSKFI--NTNMESIGKSISNNLIDKLKEASNALNTAVKGNTTGKKVSSFASGGY 2148

Qy         37 ---GLGAGSKPELGKRSAEQQKVDIQRMLQDSNMLFITGGMGGGTCTGAAPVVASVAREL 93
                |||||    |  :    | |   :|    :      |    :|      :
Db       2149 TGTGLGAGKLAFLHDKELILNKTDTANILDTVKAV-------RETAVDDSPKWGQGVKLA 2201

Qy         94 GILTVGVVSTPFRSEGPNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFR 153
              ::  |: | |      |:: | |: :  | |     | :|:: : :  ||:  :   |
Db       2202 DLIKKGITSIPSLVPNVNQSMLTNSLIPNLKK-----IEIPSKTIASSGDKTINLTNTFH 2256

Qy        154 YADDVLLEGVKGVTDL 169
               | |: |  |   :
Db       2257 I--DKLIGGESGARSM 2270


RESULT 9
US-09-134-001C-4951
; Sequence 4951, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4951
;   LENGTH: 388
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4951

  Query Match             8.9%;  Score 79;  DB 4;  Length 388;
  Best Local Similarity   23.5%;  Pred. No. 0.75;
  Matches   32;  Conservative   32;  Mismatches   48;  Indels   24;  Gaps    6;

Qy         11 VANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDSNMLFI 70
              :|||: | | :      ::|| : :    |  |       :: :  |:: |   : : :
Db         24 IANTNGQYLSKLFN----SIGKSVVEHTVIGDNP-------QRLEYVIKQCLSRFDTIVL 72

Qy         71 TGGMGGGTCTGAAPVVASVARELG--ILT----VGVVSTPFRSEGPNRTRLANAGVKELA 124
              |||:|    |       :||: ||  ::|    :  :   |: :| : |    :  |: |
Db         73 TGGLGP---TKDDLTKHTVAKVLGKNLVTDEASLNFIKNYFKEQGQDMT----SNNKQQA 125

Qy        125 KYVDTLIVVPNQNLLA 140
                ::  ||:||:| :|
Db        126 LVIEDAIVLPNKNGMA 141


RESULT 10
US-09-189-527-11
; Sequence 11, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
;  APPLICANT:  Jerome B. Posner
;  APPLICANT:  Josep O. Dalmau
;  APPLICANT:  Myrna R. Rosenfeld
;  TITLE OF INVENTION:  Ma Family Polypeptides and Anti-Ma
;  TITLE OF INVENTION:  Antibodies
;  FILE REFERENCE:  SLK98-01
;  CURRENT APPLICATION NUMBER: US/09/189,527A
;  CURRENT FILING DATE:  1998-11-10
;  NUMBER OF SEQ ID NOS:  14
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 11
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-09-189-527-11

  Query Match             8.7%;  Score 77;  DB 4;  Length 271;
  Best Local Similarity   29.1%;  Pred. No. 0.76;
  Matches   30;  Conservative    9;  Mismatches   30;  Indels   34;  Gaps    5;

Qy         37 GLGAGSKPELGKRSAEQQKVDIQRMLQDSNMLFITG---GMGGGTCTGA-----APVVAS 88
              |||  | |  | | :    : : |         |||    | | |||  |      |  ||
Db         32 GLGYVSSPGAGGRVSSHSALHLTR---------ITGPFVGTGNGTCASAPATHEIPETAS 82

Qy         89 VARE---------LGILTVGVVSTPFRSEGPNRTRLANAGVKE 122
              : ||         ||::  |          | ::   :| | ||
Db         83 ILRECCPSPRGRVLGLVGTG--------HGDSQRVASNRGRKE 117


RESULT 11
US-09-134-001C-3034
; Sequence 3034, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3034
;   LENGTH: 422
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3034

  Query Match             8.6%;  Score 76.5;  DB 4;  Length 422;
  Best Local Similarity   23.0%;  Pred. No. 1.7;
  Matches   42;  Conservative   25;  Mismatches   61;  Indels   55;  Gaps    8;

Qy         34 ITKGLGAGSKPELGKRSAEQQKVDIQRMLQDSNMLFITGGMGGGTCTGAAPVVASVAREL 93
              | :|  |   |: |::   |: : |      :|:||| ||   |       :   : | |
Db        214 ILEGTTASVPPQGGRKHPNQELIQI----DTTNILFILGGAFDG-------IDEVIKRRL 262

Qy         94 GILTVGVVSTPFRSEGPNRTRLANAGVKELAKY--------------VDTLIVVPNQNL 138
              |   :|  |     :       |     ::|  |               ::|| |   :|:
Db        263 GEKVIGFASNE-ADKYDEEALLEQIRPEDLQSYGLIPEFIGRVPIVANLETLDVAALKNI 321
```

```
Qy        139 L-----ALADKSTTMLEAFRYADDVLLE-----------------GVKGVTDLIVRPGL 175
              |     ||  : | |||     |:| ||                  | :|:   |:   |
Db        322 LTQPKNALVKQYTKMLE----LDNVELEFSEEALSAISEKAIERKTGARGLRS-IIEEAL 376

Qy        176 INL 178
              |::
Db        377 IDI 379


RESULT 12
US-09-134-001C-3477
; Sequence 3477, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3477
;   LENGTH: 327
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3477

  Query Match             8.4%;  Score 74;  DB 4;  Length 327;
  Best Local Similarity   27.7%;  Pred. No. 2.2;
  Matches   28;  Conservative   20;  Mismatches   33;  Indels   20;  Gaps    6;

Qy         95 ILTVGVVSTPFRSEGPNRTRLANAGVKELAKY----VDTLIVV---PNQNLLALADKSTT 147
              : |:|: ||  :|:      | |  : |:  |    |  ::::   ||:|   : :|  |
Db        188 VKTIGLRSTRIQSDTGEIYTLPNGMISEIVNYSATDVSPIVMIPISPNENYKVIEEKLLT 247

Qy        148 ML-------EAFRYADDVL-LEGVKG---VTDLI--VRPGL 175
                |       : |  | |:| |: | |    |  |:  |:||:
Db        248 FLPTLKNKYDIFVSAPDLLGLDSVDGNEMVIKLLAHVKPGM 288


RESULT 13
US-09-134-001C-3945
; Sequence 3945, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3945
;   LENGTH: 1529
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3945

  Query Match             8.4%;  Score 74;  DB 4;  Length 1529;
  Best Local Similarity   22.9%;  Pred. No. 21;
  Matches   41;  Conservative   29;  Mismatches   77;  Indels   32;  Gaps    7;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGL----GAGSKPELGKRSAEQQKVDIQ 59
              || || :|  ||       | |:       | : :    |  :| |: :     :  |:

Db       1182 LETVEELVGRTDLLQRSTQLKPNSKAASLQIERLIEQFDGVNTK-EISQNHHLDEGFDLN 1240

Qy         60 RMLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEG-PNRTRLA-- 116
               :  |:       |    |       || :  |::|::|   ::  :  || |  | ||
Db       1241 YLYPDARYSIENGHSFTGNY-----VVNNEQRDVGVITGSAIAKQYGEEGLPEDTILAYT 1295

Qy        117 --NAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRP 173
                :||                |:| | | :  | :     |:| : :|: | | ::  |
Db       1296 EGHAG---------------QSLAAYAPRGLT-IHHTGDANDYVGKGLSGGTVIVNAP 1337


RESULT 14
US-08-861-464-2
; Sequence 2, Application US/08861464
; Patent No. 5874210
;  GENERAL INFORMATION:
;    APPLICANT:  Guarente, Leonard P.
;    APPLICANT:  Austriaco Jr., Nicanor
;    APPLICANT:  Kennedy, Brian
;    TITLE OF INVENTION:  Genes Determining Cellular Senescence
;    TITLE OF INVENTION:  in Yeast
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/861,464
;      FILING DATE:  22-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/396,001
;      FILING DATE:  28-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/09351
;      FILING DATE:  15-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/107,408
;      FILING DATE:  16-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-6408A2Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  781-861-6240
;      TELEFAX:  781-861-9540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  450 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-861-464-2

  Query Match             8.3%;  Score 73.5;  DB 2;  Length 450;
  Best Local Similarity   23.8%;  Pred. No. 4.1;
  Matches   31;  Conservative   20;  Mismatches   46;  Indels   33;  Gaps    5;

Qy         12 ANTDCQAL-----GRSLAPHKIT-----LGKD--------ITKGLGAGSKPEL------- 46
              :|||  :|     | : | :|::       | |        :   :||||   :
Db        278 SNTDTNSLCVEGQGSAQAVNKVSGSIAICGTDYPGSENMVVPTVVGAGSSQPINVIKEDS 337

Qy         47 -----GKRSAEQQKVDIQRMLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVV 101
                  ||::: |  |:   :  :   :: | | | |     ||||       ||  : ::
Db        338 YYQWQGKKTSAQYYVNNAGVSVEDGCIWGTEGSGVGNW---APVVLGAGYTDGITYLSII 394
```

```
Qy        102 STPFRSEGPN 111
              |   | ||
Db        395 PNPNNKEAPN 404


RESULT 15
US-08-396-001-2
; Sequence 2, Application US/08396001
; Patent No. 5919618
;  GENERAL INFORMATION:
;    APPLICANT:  Guarente, Leonard P.
;    APPLICANT:  Austriaco Jr., Nicanor
;    APPLICANT:  Claus, James
;    APPLICANT:  Cole, Francesca
;    APPLICANT:  Kennedy, Brian
;    TITLE OF INVENTION:  Genes Determining Cellular Senesence in
;    TITLE OF INVENTION:  Yeast
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/396,001
;      FILING DATE:  28-FEB-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-6408A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  450 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-396-001-2

  Query Match             8.3%;  Score 73.5;  DB 2;  Length 450;
  Best Local Similarity   23.8%;  Pred. No. 4.1;
  Matches   31;  Conservative   20;  Mismatches   46;  Indels   33;  Gaps    5;

Qy         12 ANTDCQAL-----GRSLAPHKIT-----LGKD--------ITKGLGAGSKPEL------- 46
              :|||  :|     | : | :|::      | |        :   :||||   :
Db        278 SNTDTNSLCVEGQGSAQAVNKVSGSIAICGTDYPGSENMVVPTVVGAGSSQPINVIKEDS 337

Qy         47 -----GKRSAEQQKVDIQRMLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVV 101
                  ||::: |  |:   :  :   :: | | | |     ||||       ||  : ::
Db        338 YYQWQGKKTSAQYYVNNAGVSVEDGCIWGTEGSGVGNW---APVVLGAGYTDGITYLSII 394

Qy        102 STPFRSEGPN 111
              |   | ||
Db        395 PNPNNKEAPN 404


Search completed: June  2, 2003, 08:47:47
Job time : 28 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2003, 05:46:58 ; Search time 65 Seconds
(without alignments)
364.902 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	510	57.6	391	23	ABB48477	Listeria monocytog
2	495	55.9	410	22	AAU35046	Enterococcus faeca
3	495	55.9	411	22	AAU33413	Enterococcus faeca
4	493	55.7	416	23	ABB55218	Lactococcus lactis
5	485	54.8	379	22	AAG81170	Mycobacterium tube
6	480	54.2	390	21	AAY91108	Staphylococcus aur
7	480	54.2	390	21	AAY44710	S. pneumoniae mono
8	480	54.2	390	22	AAU34274	Staphylococcus aur
9	480	54.2	392	22	AAU37130	Staphylococcus aur
10	480	54.2	394	23	ABP39989	Staphylococcus epi
11	472	53.3	335	19	AAW80614	S.pneumoniae filam
12	472	53.3	419	20	AAW99120	Streptococcus pneu
13	472	53.3	419	21	AAY44711	Monomeric filament
14	472	53.3	419	22	AAU37886	Streptococcus pneu
15	470	53.1	442	22	AAG92106	C glutamicum prote
16	466	52.7	421	22	AAU60468	Propionibacterium
17	464	52.4	394	22	AAU36414	Pseudomonas aerugi
18	464	52.4	439	23	ABP25525	Streptococcus poly
19	463	52.3	397	21	AAB01328	FtsZ2 polypeptide
20	463	52.3	437	19	AAW41733	Arabidopsis chloro
21	456.5	51.6	427	23	ABP25524	Streptococcus poly
22	452	51.1	383	21	AAB15908	E. coli proliferat
23	452	51.1	383	21	AAY91110	Escherichia coli m
24	452	51.1	383	22	AAU34440	E. coli cellular p
25	452	51.1	413	21	AAB01330	xanFtsZ1 polypepti
26	447	50.5	383	23	AAU76922	Wild-type E. coli
27	447	50.5	383	23	AAU76923	E. coli ftsz prote
28	447	50.5	383	23	AAU76924	E. coli ftsz prote
29	447	50.5	383	23	AAU76925	E. coli ftsz prote
30	447	50.5	421	22	AAU35599	Haemophilus influe
31	441	49.8	374	19	AAW70503	Brevibacterium fla
32	441	49.8	374	19	AAW41787	Brevibacterium fla
33	440	49.7	361	19	AAW51423	Neisseria meningit
34	439	49.6	325	21	AAB01332	FtsZ1 polypeptide
35	435	49.2	363	21	AAG46656	Arabidopsis thalia
36	435	49.2	372	21	AAG46655	Arabidopsis thalia
37	435	49.2	433	21	AAG46654	Arabidopsis thalia
38	435	49.2	433	21	AAB01327	FtsZ1 polypeptide
39	433	48.9	411	21	AAB01329	BnFtsZ1 polypeptid
40	429	48.5	433	19	AAW41732	Arabidopsis chloro
41	424.5	48.0	395	22	AAB96802	Putative P. abyssi
42	379	42.8	385	22	AAU35960	Helicobacter pylor
43	323.5	36.6	413	22	AAB96344	Putative P. abyssi
44	304	34.4	357	21	AAB01333	FtsZ1 polypeptide
45	285	32.2	267	21	AAG04705	Arabidopsis thalia

ALIGNMENTS

RESULT 1
ABB48477
ID ABB48477 standard; Protein; 391 AA.
XX
AC ABB48477;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #1181.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;

```
XX
DR   WPI; 2002-010914/01.
XX
PT   Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT   and prevention of Listeria and related bacterial infections, and
PT   related polypeptides -
XX
PS   Claim 6; SEQ ID No 1182; 192pp; French.
XX
CC   The present invention relates to the genome sequence of Listeria
CC   monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC   it are useful for selecting probes and primers for detecting genes in L.
CC   monocytogenes and related organisms, and for studying genetic
CC   polymorphisms and other genomes. The present sequence is a protein
CC   encoded by the genome sequence of the present invention. Proteins
CC   expressed from the genome sequence are useful for raising specific
CC   antibodies, identification of L. monocytogenes and related organisms, and
CC   for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC   B12. The genome sequence and proteins encoded by it are also useful for
CC   selecting compounds that regulate gene expression and cell replication
CC   and modulate L. monocytogenes-related diseases. In addition, the genome
CC   sequence and proteins encoded by it are useful in pharmaceutical and
CC   vaccines compositions for the treatment or prevention of infections by L.
CC   monocytogenes and related organisms.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   391 AA;

  Query Match             57.6%;  Score 510;  DB 23;  Length 391;
  Best Local Similarity   56.0%;  Pred. No. 1.5e-45;
  Matches   98;  Conservative   35;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  ||| |||  : |  |: :|  :|:||||||: ||:||::||: :  |:  |:
Db         35 VQGVEFISVNTDAQALNLAKAETKLQIGTKLTRGLGAGAVPEIGKKAAEESREQIEEALK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              |:|:|:| |||||| ||||||:| :|:|:| ||||||: ||  ||| ||: |  | : :
Db         95 GSDMVFVTAGMGGGTGTGAAPVIAQIAKEMGALTVGVVTRPFGFEGPKRTKQALTGTEAM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               : ||||||:||  || : ||:| |||||| ||:|| :||:|::|||  ||||||
Db        155 KEAVDTLIVIPNDRLLQIVDKNTPMLEAFREADNVLRQGVQGISDLIAVPGLINL 209


RESULT 2
AAU35046
ID   AAU35046 standard; Protein; 410 AA.
XX
AC   AAU35046;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Enterococcus faecalis cellular proliferation protein #333.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Enterococcus faecalis.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS52905.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 10639; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   410 AA;

  Query Match             55.9%;  Score 495;  DB 22;  Length 410;
  Best Local Similarity   57.1%;  Pred. No. 6.2e-44;
  Matches  100;  Conservative   28;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  |||| |||  | |   | ||   |:||||||:||:|:::||: :  |   ||
Db         36 VKGVEFITANTDVQALKHSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQVISESLQ 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ::|:|||  |||||| |||||||| :|:||| ||||||: ||  ||| | | |  |:  |
Db         96 GADMIFITAGMGGGTGTGAAPVVAKIAKELGALTVGVVTRPFSFEGPKRGRFAAEGIALL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        156 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 210


RESULT 3
AAU33413
ID   AAU33413 standard; Protein; 411 AA.
XX
AC   AAU33413;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Enterococcus faecalis cellular proliferation protein #49.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Enterococcus faecalis.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
```

XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;.
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS51272.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 4909; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 411 AA;

Query Match 55.9%; Score 495; DB 22; Length 411;
Best Local Similarity 57.1%; Pred. No. 6.3e-44;
Matches 100; Conservative 28; Mismatches 47; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||| |||| |||  | |   | ||   |:||||||:||:|:::||: :  |   ||
Db         36 VKGVEFITANTDVQALKHSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQVISESLQ 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:||| |||||| |||||||| :|:||| |||||: ||  ||| | | |  |:  |
Db         96 GADMIFITAGMGGGTGTGAAPVVAKIAKELGALTVGVVTRPFSFEGPKRGRFAAEGIALL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        156 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 210
```

RESULT 4
ABB55218
ID ABB55218 standard; Protein; 416 AA.
XX
AC ABB55218;
XX
DT 16-MAY-2002 (first entry)
XX
DE Lactococcus lactis protein ftsZ.
XX
KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
OS Lactococcus lactis IL1403.
XX
PN FR2807446-A1.
XX
PD 12-OCT-2001.
XX
PF 11-APR-2000; 2000FR-0004630.
XX
PR 11-APR-2000; 2000FR-0004630.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX
PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX
DR WPI; 2002-043418/06.
XX
PT New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species -
XX
PS Claim 6; SEQ ID No 1920; 2504pp; French.
XX
CC The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 416 AA;

Query Match 55.7%; Score 493; DB 23; Length 416;
Best Local Similarity 56.0%; Pred. No. 1e-43;
Matches 98; Conservative 29; Mismatches 48; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              : ||||| |||| |||  | |   | ||  :|:||||||::||:|||:|||:    : : |:
Db         36 VSGVEFIAANTDVQALRSSKADTVIQLGPKLTRGLGAGAQPEVGKRAAEESAETVSQALE 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               |:|:||| |||||| |||||||:| :|:||| |||||| ||  ||  |:  |  |:: |
Db         96 GSDMIFITAGMGGGTGTGAAPVIAQIAKELGALTVGVVTRPFGFEGSKRSYFATEGIEAL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                |||::: | ||| : || | : || | ||:|| :||:|||||  ||:|||
Db        156 RANVDTLLIISNNNLLEIVDKKTPLTEALREADNVLRQGVQGVTDLITNPGMINL 210
```

RESULT 5
AAG81170
ID AAG81170 standard; Protein; 379 AA.
XX
AC AAG81170;
XX
DT 04-SEP-2001 (first entry)
XX
DE Mycobacterium tuberculosis potential drug target protein SEQ ID 221.
XX
KW Drug target; growth; organism viability; characterisation.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200135317-A1.
XX
PD 17-MAY-2001.
XX
PF 13-NOV-2000; 2000WO-US31152.

XX
PR 12-NOV-1999; 99US-0165086.
PR 12-NOV-1999; 99US-0165124.
PR 01-FEB-2000; 2000US-0179531.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Eisenberg D, Rotstein SH, Marcotte EM;
XX
DR WPI; 2001-329193/34.
DR N-PSDB; AAH52021.
XX
PT Identifying nucleotide or polypeptide sequence for use as drug target,
PT involves providing algorithm that analyzes a functional relationship
PT between nucleotide or polypeptide sequences, and comparing the
PT sequences -
XX
PS Disclosure; Page 172; 207pp; English.
XX
CC This invention relates to a method for identifying a nucleotide or
CC polypeptide sequence that may be a drug target, or essential for growth
CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
CC represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium
CC tuberculosis proteins which are potential drug targets. The DNA and
CC protein sequences are used to illustrate the method of the invention. The
CC method involves providing an unknown nucleotide or polypeptide sequences,
CC and comparing it to a number of sequences along with at least one
CC algorithm capable of analysing a functional relationship between
CC nucleotide and polypeptide sequences. The method is useful for
CC characterising the function of nucleic acids and polypeptides that may be
CC useful as a target for a drug or essential for the growth or viability of
CC an organism.
XX
SQ Sequence 379 AA;

Query Match 54.8%; Score 485; DB 22; Length 379;
Best Local Similarity 54.3%; Pred. No. 6.5e-43;
Matches 95; Conservative 35; Mismatches 45; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||||  ||| ||| | |  |: :|:| |:|||||: ||:|:::||  | :|: :|:
Db         32 LKGVEFIAINTDAQALLMSDADVKLDVGRDSTRGLGAGADPEVGRKAAEDAKDEIEELLR 91

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| | |||| || ||||||:||:|| ||||||: ||  ||  |:  |  |:  |
Db         92 GADMVFVTAGEGGGTGTGGAPVVASIARKLGALTVGVVTRPFSFEGKRRSNQAENGIAAL 151

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                :  |||||:||  || : | : :::::||| ||:||| ||:|:|||| |||||:
Db        152 RESCDTLIVIPNDRLLQMGDAAVSLMDAFRSADEVLLNGVQGITDLITTPGLINV 206
```

RESULT 6
AAY91108
ID AAY91108 standard; Protein; 390 AA.
XX
AC AAY91108;
XX
DT 15-SEP-2000 (first entry)
XX
DE Staphylococcus aureus monomeric FtsZ protein sequence SEQ ID NO:2.
XX
KW Staphylococcus aureus; monomeric FtsZ; monomeric FtsA; antibacterial;
KW Escherichia coli; multimeric protein; bacterial infection; vaccine;
KW microbial infection; screening.
XX
OS Staphylococcus aureus.
XX
PN WO200024252-A1.
XX
PD 04-MAY-2000.
XX
PF 20-OCT-1999; 99WO-US24653.
XX
PR 23-OCT-1998; 98US-0105315.
PR 19-APR-1999; 99US-0129965.
PR 03-MAY-1999; 99US-0132333.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Yan K, Pearce KH;
XX
DR WPI; 2000-350517/30.
DR N-PSDB; AAA39395.
XX
PT Multimeric FtsZ:FtsA polypeptide, useful as a vaccine for treating
PT bacterial infections and for diagnosing diseases associated with
PT bacterial infection -
XX
PS Claim 1; Page 2; 45pp; English.
XX
CC The present invention describes isolated multimeric FtsZ:FtsA
CC polypeptides (I) comprising the protein sequences given in AAY91108,
CC AAY91109, AAY91110 and AAY91111. Also described are: (1) an antibody
CC (Ab) specific for (I); (2) a method (M1) for treating an individual
CC requiring enhanced or reduced activity or expression of (I) comprising
CC administering an agonist or antagonist to (I), respectively; (3) a
CC method (M2) for diagnosing a disease or susceptibility to a disease
CC related to expression or activity of (I) comprising determining the
CC presence or absence of a mutation in the nucleotide sequence encoding
CC (I) and/or analysing for the presence or quantity of (I) in a sample;
CC and (4) a method (M3) for screening to identify compounds that activate
CC or inhibit the function of (I) selected from: (a) measuring the binding
CC of a candidate compound to the polypeptide using a label; (b) measuring
CC the binding of a candidate compound to (I) in the presence of a labeled
CC competitor; (c) testing whether the candidate compound results in a
CC signal generated by activation or inhibition of (I); or (d) detecting
CC the effect of a candidate compound on the production of mRNA encoding
CC (I) using e.g. enzyme linked immunosorbent assay (ELISA). AAY91108 to
CC AAY91111 are useful as vaccines for treating microbial infections such
CC as Staphylococcus, Streptococcus and Escherichia coli infections. The
CC polypeptides are also useful for screening for antibacterial compounds
CC and for detecting diseases associated with microbial infections. The
CC present sequence represents Staphylococcus monomeric FtsZ.
XX
SQ Sequence 390 AA;

Query Match 54.2%; Score 480; DB 21; Length 390;
Best Local Similarity 54.3%; Pred. No. 2.3e-42;
Matches 95; Conservative 33; Mismatches 47; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              :  ||||  ||| |||  | |  || :|: :|:|||||: ||:||::||: :  |:  :|
Db         35 MNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQ 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| |||||| |||||||| :|:|:| ||||||: ||  ||  |   | |||: :
Db         95 GADMVFVTSGMGGGTGTGAAPVVAKIAKEMGALTVGVVTRPFSFEGRKRQTQAAAGVEAM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                |||||:||  || : |||| |:|||: ||:|| :||:|::|||   | :||
Db        155 KAAVDTLIVIPNDRLLDIVDKSTPMMEAFKEADNVLRQGVQGISDLIAVSGEVNL 209
```

RESULT 7
AAY44710
ID AAY44710 standard; Protein; 390 AA.
XX
AC AAY44710;
XX
DT 25-APR-2000 (first entry)
XX
DE S. pneumoniae monomeric filamentation temperature sensitive protein.
XX
KW Monomeric filamentation temperature sensitive protein;
KW FtsZ; dental treatment; vaccine; wound infection; gastric ulcer;

KW gastrointestinal cancer; Helicobacter pylori infection; gastritis;
KW antibacterial; antiulcer; anti-inflammatory.
XX
OS Streptococcus pneumoniae.
XX
PN WO200001801-A1.
XX
PD 13-JAN-2000.
XX
PF 30-JUN-1999; 99WO-US14890.
XX
PR 02-JUL-1998; 98US-0091680.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Hensley CP, Pearce KH, Sossong TM;
XX
DR WPI; 2000-171008/15.
DR N-PSDB; AAZ49958.
XX
PT Isolated polypeptide useful as a vaccine and for screening
PT antimicrobial compounds for treating microbial diseases -
XX
PS Claim 1; Page 3; 52pp; English.
XX
CC The present sequence is the S. pneumoniae monomeric filamentation
CC temperature sensitive protein (FtsZ). This is related by amino acid
CC sequence homology to Lactococcus lactis FtsZ AAU74322. Multimeric FtsZ
CC protein is useful in identifying compounds that modulate its function.
CC The protein and its variants are useful in diagnosing and treating
CC microbial diseases and for prophylactic use in dental treatment. They
CC can be used as vaccines to prevent wound infections by bacteria
CC especially S. pneumoniae. Agonists and antagonists are useful in
CC treating diseases caused by Helicobacter pylori infections e.g.
CC gastrointestinal cancer, gastric ulcer and gastritis.
XX
SQ Sequence 390 AA;

Query Match 54.2%; Score 480; DB 21; Length 390;
Best Local Similarity 54.3%; Pred. No. 2.3e-42;
Matches 95; Conservative 33; Mismatches 47; Indels 0; Gaps 0;

Qy 4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
: |||| ||| ||| | | || :|: :|:|||||: ||:||::||: : |: :|
Db 35 MNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQ 94

Qy 64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
::|:|:| |||||| |||||||| :|:|:| ||||||: || || | | |||: :
Db 95 GADMVFVTSGMGGGTGTGAAPVVAKIAKEMGALTVGVVTRPFSFEGRKRQTQAAAGVEAM 154

Qy 124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
||||||:|| || : |||| |:|||: ||:|| :||:|::||| | :||
Db 155 KAAVDTLIVIPNDRLLDIVDKSTPMMEAFKEADNVLRQGVQGISDLIAVSGEVNL 209

RESULT 8
AAU34274
ID AAU34274 standard; Protein; 390 AA.
XX
AC AAU34274;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #550.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS52133.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 5770; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 390 AA;

Query Match 54.2%; Score 480; DB 22; Length 390;
Best Local Similarity 54.3%; Pred. No. 2.3e-42;
Matches 95; Conservative 33; Mismatches 47; Indels 0; Gaps 0;

Qy 4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
: |||| ||| ||| | | || :|: :|:|||||: ||:||::||: : |: :|
Db 35 MNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQ 94

Qy 64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
::|:|:| |||||| |||||||| :|:|:| |||||: || || | | |||: :
Db 95 GADMVFVTSGMGGGTGTGAAPVVAKIAKEMGALTVGVVTRPFSFEGRKRQTQAAAGVEAM 154

Qy 124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
||||||:|| || : |||| |:|||: ||:|| :||:|::||| | :||
Db 155 KAAVDTLIVIPNDRLLDIVDKSTPMMEAFKEADNVLRQGVQGISDLIAVSGEVNL 209

RESULT 9
AAU37130
ID AAU37130 standard; Protein; 392 AA.
XX
AC AAU37130;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1300.
XX

```
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS54989.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 12723; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   392 AA;

  Query Match             54.2%;  Score 480;  DB 22;  Length 392;
  Best Local Similarity   54.3%;  Pred. No. 2.3e-42;
  Matches   95;  Conservative   33;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              :  ||||  ||| |||  | |  || :|: :|:|||||: ||:||::||: :  |:  :|
Db         35 MNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQ 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| |||||| |||||||| :|:|:| ||||||: ||  ||  |   | |||: :
Db         95 GADMVFVTSGMGGGTGTGAAPVVAKIAKEMGALTVGVVTRPFSFEGRKRQTQAAAGVEAM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                  ||||||:||  || : |||| |:|||: ||:|| :|||:|::|||    | :||
Db        155 KAAVDTLIVIPNDRLLDIVDKSTPMMEAFKEADNVLRQGVQGISDLIAVSGEVNL 209


RESULT 10
ABP39989
ID   ABP39989 standard; Protein; 394 AA.
XX
AC   ABP39989;
XX
DT   24-JUL-2002  (first entry)
XX
DE   Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4834.
XX
KW   Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW   antibacterial; gene therapy.
XX
OS   Staphylococcus epidermidis.
XX
PN   US6380370-B1.
XX
PD   30-APR-2002.
XX
PF   13-AUG-1998;   98US-0134001.
XX
PR   14-AUG-1997;   97US-055779P.
PR   08-NOV-1997;   97US-064964P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm LA,  Bush D;
XX
DR   WPI; 2002-381255/41.
DR   N-PSDB; ABN92534.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT   polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS   Disclosure; SEQ ID 4834; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences
CC   can also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life
CC   cycle or inhibit S. epidermidis infection.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   USPTO web site.
XX
SQ   Sequence   394 AA;

  Query Match             54.2%;  Score 480;  DB 23;  Length 394;
  Best Local Similarity   54.3%;  Pred. No. 2.3e-42;
  Matches   95;  Conservative   33;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              :  ||||  ||| |||  | |  || :|: :|:|||||: ||:||::||: :  |:  :|
Db         35 MNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQ 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| |||||| |||||||| :|:|:| ||||||: ||  ||  |   | |||: :
Db         95 GADMVFVTAGMGGGTGTGAAPVVAKIAKEMGALTVGVVTRPFGFEGRKRQTQAAAGVESM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                  ||||||:||  || : |||| |:|||: ||:|| :|||:|::|||    | :||
Db        155 KAAVDTLIVIPNDRLLDIVDKSTPMMEAFKEADNVLRQGVQGISDLIAVSGEVNL 209


RESULT 11
AAW80614
ID   AAW80614 standard; Protein; 335 AA.
XX
AC   AAW80614;
XX
DT   24-DEC-1998  (first entry)
XX
```

```
DE   S.pneumoniae filamentous temperature sensitive, cell septation protein.
XX
KW   Streptococcus pneumoniae protein; recombinant; gene expression;
KW   DNA chip; virulence; antibody; infection; detection; treatment.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO9826072-A1.
XX
PD   18-JUN-1998.
XX
PF   09-DEC-1997;   97WO-US22578.
XX
PR   13-DEC-1996;   96US-0036281.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Baltz RH,  Burgett SG,  Dehoff BS,  Hoskins JA,  Jaskunas SR;
PI   Mills BJ,  Norris FH,  Peery RB,  Rockey PK,  Rosteck PR;
PI   Skatrud PL,  Smith MC,  Solenberg PJ,  Treadway PJ;
PI   Young Bellido ML;
XX
DR   WPI; 1998-348529/30.
DR   N-PSDB; AAV65296.
XX
PT   Streptococcus pneumoniae nucleic acid sequences - used in DNA chips
PT   for evaluating gene expression, and identification of virulence
PT   genes
XX
PS   Claim 3; Pages 182-183; 333pp; English.
XX
CC   This sequence represents a S. pneumoniae filamentous temperature
CC   sensitive, cell septation protein. The invention provides DNA sequences
CC   (AAV65201 to AAV65304) from the S. pneumoniae genome and corresponding
CC   protein sequences(AAW80605 to AAW80728). A recombinant host containing a
CC   vector comprising any of the above nucleic acids can be used for the
CC   recombinant expression of the protein sequences. The invention also
CC   provides a DNA chip having arrayed on it at least 15 base pair fragment
CC   of any one or more of these DNA sequences. The DNA chip can be used
CC   methods for evaluating gene expression in S. pneumoniae and for
CC   identifying virulence genes in S. pneumoniae. Antibodies that
CC   selectively bind to the above proteins or peptide fragments can be used
CC   to treat S. pneumoniae infection. The antibodies can also be used to
CC   detect S. pneumoniae cells.
XX
SQ   Sequence   335 AA;

  Query Match             53.3%;  Score 472;  DB 19;  Length 335;
  Best Local Similarity   52.6%;  Pred. No. 1.3e-41;
  Matches   91;  Conservative   35;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          6 GVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDS 65
              ||||| |||| ||| : |   | ||  :|:||||| :||:|:::||: :  :   :  :
Db         38 GVEFIAANTDVQALSSTKAETVIQLGPKLTRGLGAGGQPEVGRKAAEESEETLTEAISGA 97

Qy         66 NMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAK 125
              :|:||| |||||: ||||||:| :|::|| |||||: ||  ||  | : |  |: :| :
Db         98 DMVFITAGMGGGSGTGAAPVIARIAKDLGALTVGVVTRPFGFEGSKRGQFAVEGINQLRE 157

Qy        126 YVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :||||::: | ||| : || | :|||   ||:|| :||:|:|||| ||||||
Db        158 HVDTLLIISNNNLLEIVDKKTPLLEALSEADNVLRQGVQGITDLITNPGLINL 210


RESULT 12
AAW99120
ID   AAW99120 standard; Protein; 419 AA.
XX
AC   AAW99120;
XX
DT   17-MAY-1999  (first entry)
XX
DE   Streptococcus pneumoniae ftsZ protein.
XX
KW   Streptococcus pneumoniae; ftsZ; antibacterial; microbial disease;
KW   otitis media; bacteraemia; conjunctivitis; pneumonia; sinusitis;
KW   pleural empyema; endocarditis; meningitis.
XX
OS   Streptococcus pneumoniae.
XX
PN   EP899334-A2.
XX
PD   03-MAR-1999.
XX
PF   30-JUL-1998;   98EP-0306077.
XX
PR   12-AUG-1997;   97US-0055720.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Fueyo JL,  Lonetto MA;
XX
DR   WPI; 1999-144805/13.
DR   N-PSDB; AAX19269.
XX
PT   New Streptococcus pneumoniae ftsZ polypeptide and polynucleotide -
PT   useful as diagnostic reagents and for prevention and treatment of
PT   Streptococci infections which cause bacteraemia, otitis media and
PT   meningitis
XX
PS   Claim 1; Page 4; 37pp; English.
XX
CC   The present sequence is ftsZ polypeptide isolated from Streptococcus
CC   pneumoniae. FtsZ polynucleotides and polypeptides are useful for
CC   diagnosing susceptibility to diseases by detecting mutations or
CC   polymorphisms in the ftsZ gene or analysing for the presence of amount of
CC   ftsZ polypeptide expressed in a patient sample. FtsZ PCR probes are
CC   useful for diagnosing diseases, and can characterise the response of
CC   the infectious organism to drugs. FtsZ polypeptides and polynucleotides
CC   are also useful for screening for antagonists, agonists and drugs against
CC   infectious micro-organisms. FtsZ agonists and antagonists are
CC   bacteriostatic and bacteriocidal compounds which can be used in treatment
CC   to enhance (agonist) or block (antagonist or antisense sequence) ftsZ
CC   activity, therefore treating microbial diseases, especially Streptococcus
CC   pneumoniae diseases including otitis media, bacteraemia, conjunctivitis,
CC   pneumonia, sinusitis, pleural empyema, endocarditis and especially
CC   meningitis. Epitopes of ftsZ polypeptides and polynucleotides are useful
CC   immunogens for producing anti-ftsZ antibodies for prevention of bacterial
CC   infections, and ftsZ polynucleotides can be used in genetic immunisation
CC   (gene therapy) to prevent infections. FtsZ polypeptides, polynucleotides
CC   and their (ant)agonists can prevent adhesion of bacteria to matrix
CC   proteins, and are useful for use on wounds and body implants to prevent
CC   bacterial infection. FtsZ polypeptides and polynucleotides may also be
CC   used as reagents for differential screening methods e.g. using ftsZ
CC   probes in RT-PCR to identify and quantify genes expressed in bacterial
CC   tissue.
XX
SQ   Sequence   419 AA;

  Query Match             53.3%;  Score 472;  DB 20;  Length 419;
  Best Local Similarity   52.6%;  Pred. No. 1.8e-41;
  Matches   91;  Conservative   35;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          6 GVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDS 65
              ||||| |||| ||| : |   | ||  :|:||||| :||:|:::||: :  :   :  :
Db         38 GVEFIAANTDVQALSSTKAETVIQLGPKLTRGLGAGGQPEVGRKAAEESEETLTEAISGA 97

Qy         66 NMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAK 125
              :|:||| |||||: ||||||:| :|::|| |||||: ||  ||  | : |  |: :| :
Db         98 DMVFITAGMGGGSGTGAAPVIARIAKDLGALTVGVVTRPFGFEGSKRGQFAVEGINQLRE 157

Qy        126 YVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :||||::: | ||| : || | :|||   ||:|| :||:|:|||| ||||||
Db        158 HVDTLLIISNNNLLEIVDKKTPLLEALSEADNVLRQGVQGITDLITNPGLINL 210
```

```
RESULT 13
AAY44711
ID   AAY44711 standard; Protein; 419 AA.
XX
AC   AAY44711;
XX
DT   25-APR-2000  (first entry)
XX
DE   Monomeric filamentation temperature sensitive protein.
XX
KW   Monomeric filamentation temperature sensitive protein;
KW   FtsZ; dental treatment; vaccine; wound infection; gastric ulcer;
KW   gastrointestinal cancer; Helicobacter pylori infection; gastritis;
KW   antibacterial; antiulcer; anti-inflammatory.
XX
OS   Staphylococcus aureus.
XX
PN   WO200001801-A1.
XX
PD   13-JAN-2000.
XX
PF   30-JUN-1999;   99WO-US14890.
XX
PR   02-JUL-1998;   98US-0091680.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Hensley CP,  Pearce KH,  Sossong TM;
XX
DR   WPI; 2000-171008/15.
DR   N-PSDB; AAZ49959.
XX
PT   Isolated polypeptide useful as a vaccine and for screening
PT   antimicrobial compounds for treating microbial diseases -
XX
PS   Claim 1; Page 4; 52pp; English.
XX
CC   The present sequence is the monomeric filamentation
CC   temperature sensitive protein (FtsZ). This is related by amino acid
CC   sequence homology to Lactococcus lactis FtsZ AAU74322. Multimeric FtsZ
CC   protein is useful in identifying compounds that modulate its function.
CC   The protein and its variants are useful in diagnosing and treating
CC   microbial diseases and for prophylactic use in dental treatment. They
CC   can be used as vaccines to prevent wound infections by bacteria
CC   especially S. pneumoniae. Agonists and antagonists are useful in
CC   treating diseases caused by Helicobacter pylori infections e.g.
CC   gastrointestinal cancer, gastric ulcer and gastritis.
CC   Note: The specification is unclear as to whether this sequence was
CC   isolated from Streptococcus pneumoniae or Staphylococcus aureus.
XX
SQ   Sequence   419 AA;

  Query Match             53.3%;  Score 472;  DB 21;  Length 419;
  Best Local Similarity   52.6%;  Pred. No. 1.8e-41;
  Matches   91;  Conservative   35;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          6 GVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDS 65
              ||||| |||| ||| : |   | ||  :|:||||| :||:|:::||: :  :   :  :
Db         38 GVEFIAANTDVQALSSTKAETVIQLGPKLTRGLGAGGQPEVGRKAAEESEETLTEAISGA 97

Qy         66 NMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAK 125
              :|:||| |||||: ||||||:| :|::|| ||||||: ||  ||  | : |  |: :| :
Db         98 DMVFITAGMGGGSGTGAAPVIARIAKDLGALTVGVVTRPFGFEGSKRGQFAVEGINQLRE 157

Qy        126 YVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :||||::: | ||| : || | :|||   ||:|| :||:|:|||| ||||||
Db        158 HVDTLLIISNNNLLEIVDKKTPLLEALSEADNVLRQGVQGITDLITNPGLINL 210


RESULT 14
AAU37886
ID   AAU37886 standard; Protein; 419 AA.
XX
AC   AAU37886;
XX
DT   14-FEB-2002  (first entry)
XX
DE   Streptococcus pneumoniae cellular proliferation protein #315.
XX
KW   Antisense; prokaryotic cellular proliferation protein;
KW   antibiotic; antibacterial; drug design.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   N-PSDB; AAS55745.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Example 3; Seq ID No 13479; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence represents an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   419 AA;

  Query Match             53.3%;  Score 472;  DB 22;  Length 419;
  Best Local Similarity   52.6%;  Pred. No. 1.8e-41;
  Matches   91;  Conservative   35;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          6 GVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDS 65
              ||||| |||| ||| : |   | ||  :|:||||| :||:|:::||: :  :   :  :
Db         38 GVEFIAANTDVQALSSTKAETVIQLGPKLTRGLGAGGQPEVGRKAAEESEETLTEAISGA 97

Qy         66 NMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAK 125
              :|:||| |||||: ||||||:| :|::|| ||||||: ||  ||  | : |  |: :| :
Db         98 DMVFITAGMGGGSGTGAAPVIARIAKDLGALTVGVVTRPFGFEGSKRGQFAVEGINQLRE 157

Qy        126 YVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
```

```
              :||||::: | ||| : || | :|||   ||:|| :||:|:|||| ||||||
Db        158 HVDTLLIISNNNLLEIVDKKTPLLEALSEADNVLRQGVQGITDLITNPGLINL 210


RESULT 15
AAG92106
ID   AAG92106 standard; Protein; 442 AA.
XX
AC   AAG92106;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum protein fragment SEQ ID NO: 5860.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW   organic acid synthesis.
XX
OS   Corynebacterium glutamicum.
XX
PN   EP1108790-A2.
XX
PD   20-JUN-2001.
XX
PF   18-DEC-2000; 2000EP-0127688.
XX
PR   16-DEC-1999;   99JP-0377484.
PR   07-APR-2000; 2000JP-0159162.
PR   03-AUG-2000; 2000JP-0280988.
XX
PA   (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI   Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI   Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR   WPI; 2001-376931/40.
DR   N-PSDB; AAH67325.
XX
PT   Novel polynucleotides derived from Coryneform bacteria, for identifying
PT   mutation point of a gene, measuring expression of a gene, analysing
PT   expression profile or pattern of a gene and identifying homologous gene
PT   -
XX
PS   Claim 17; SEQ ID NO: 5860; 246pp + Sequence Listing; English.
XX
CC   The present invention provides a number of nucleotide and protein
CC   sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC   are useful for identifying the mutation point of a gene derived from a
CC   mutant of coryneform bacterium, measuring expression amount and
CC   analysing the expression profile or expression pattern of a gene derived
CC   from Coryneform bacterium, and identifying a homologue of a gene derived
CC   from coryneform bacterium. Coryneform bacteria are useful for producing
CC   amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC   particularly L-lysine. The present sequence is a protein described
CC   in the exemplification of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   European Patent Office.
XX
SQ   Sequence   442 AA;

  Query Match             53.1%;  Score 470;  DB 22;  Length 442;
  Best Local Similarity   52.6%;  Pred. No. 3.1e-41;
  Matches   92;  Conservative   36;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:||||  ||| ||| | |  |: :|:: |:|||||: ||:|: |||  | :|:  ::
Db         32 LKGVEFIAVNTDSQALMFSDADVKLDIGREATRGLGAGANPEVGRASAEDHKNEIEETIK 91

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:|:| | |||| |||||||| :|:::| ||:|||: ||  ||  ||| |  |:  |
Db         92 GADMVFVTAGEGGGTGTGAAPVVAGIAKKMGALTIGVVTKPFEFEGRRRTRQAEEGIAAL 151

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178

              :  |||||:||  || | | : :::|||| ||:||  ||:|:|:|| ||:||:
Db        152 KEVCDTLIVIPNDRLLELGDANLSIMEAFRAADEVLHNGVQGITNLITIPGVINV 206


Search completed: June  2, 2003, 08:33:21
Job time : 72 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2003, 08:42:25 ; Search time 233 Seconds
(without alignments)
77.330 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	885	100.0	178	9	US-09-770-509-2	Sequence 2, Appli
2	527	59.5	401	9	US-09-770-509-17	Sequence 17, Appl
3	518	58.5	452	9	US-09-770-509-14	Sequence 14, Appl
4	505	57.1	508	9	US-09-770-509-15	Sequence 15, Appl
5	496	56.0	581	9	US-09-770-509-13	Sequence 13, Appl
6	495	55.9	410	10	US-09-815-242-10639	Sequence 10639, A
7	495	55.9	411	10	US-09-815-242-4909	Sequence 4909, Ap
8	490	55.4	407	9	US-09-770-509-16	Sequence 16, Appl
9	485	54.8	379	9	US-09-712-363-221	Sequence 221, App
10	482	54.5	583	9	US-09-770-509-11	Sequence 11, Appl
11	481	54.4	420	9	US-09-770-509-10	Sequence 10, Appl
12	480	54.2	390	10	US-09-815-242-5770	Sequence 5770, Ap
13	480	54.2	392	10	US-09-815-242-12723	Sequence 12723, A
14	475	53.7	590	9	US-09-770-509-12	Sequence 12, Appl
15	472	53.3	419	10	US-09-754-608-2	Sequence 2, Appli
16	472	53.3	419	10	US-09-815-242-13479	Sequence 13479, A
17	470	53.1	442	9	US-09-738-626-5860	Sequence 5860, Ap
18	468	52.9	428	9	US-09-770-509-25	Sequence 25, Appl
19	467	52.8	430	9	US-09-770-509-26	Sequence 26, Appl
20	464	52.4	394	10	US-09-815-242-12007	Sequence 12007, A
21	463	52.3	397	9	US-09-770-509-20	Sequence 20, Appl
22	462	52.2	464	9	US-09-770-509-22	Sequence 22, Appl
23	461	52.1	458	9	US-09-770-509-21	Sequence 21, Appl
24	455	51.4	398	9	US-09-770-509-23	Sequence 23, Appl
25	452	51.1	383	10	US-09-912-020-265	Sequence 265, App
26	452	51.1	383	10	US-09-815-242-10033	Sequence 10033, A
27	451	51.0	408	9	US-09-770-509-31	Sequence 31, Appl
28	451	51.0	413	9	US-09-770-509-29	Sequence 29, Appl
29	451	51.0	413	9	US-09-770-509-32	Sequence 32, Appl
30	451	51.0	419	9	US-09-770-509-30	Sequence 30, Appl
31	448	50.6	423	9	US-09-770-509-28	Sequence 28, Appl
32	447	50.5	383	9	US-09-908-504A-1	Sequence 1, Appli
33	447	50.5	421	10	US-09-815-242-11192	Sequence 11192, A
34	446	50.4	468	9	US-09-770-509-19	Sequence 19, Appl
35	440	49.7	483	9	US-09-770-509-18	Sequence 18, Appl
36	436	49.3	368	9	US-09-770-509-24	Sequence 24, Appl
37	429	48.5	433	9	US-09-770-509-27	Sequence 27, Appl
38	379	42.8	385	10	US-09-815-242-11553	Sequence 11553, A
39	271	30.6	223	10	US-09-754-608-4	Sequence 4, Appli
40	196	22.1	283	10	US-09-765-272-124	Sequence 124, App
41	160	18.1	73	9	US-09-770-509-4	Sequence 4, Appli
42	90.5	10.2	2285	10	US-09-932-183A-2	Sequence 2, Appli
43	84	9.5	448	9	US-09-738-626-5103	Sequence 5103, Ap
44	82.5	9.3	308	9	US-09-738-626-6317	Sequence 6317, Ap
45	78.5	8.9	315	10	US-09-815-242-13514	Sequence 13514, A

ALIGNMENTS

RESULT 1
US-09-770-509-2
; Sequence 2, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
; APPLICANT: Katagiri, F.
; TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES
; FILE REFERENCE: NADII.018A
; CURRENT APPLICATION NUMBER: US/09/770,509
; CURRENT FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Phytophthora infestans
US-09-770-509-2

Query Match 100.0%; Score 885; DB 9; Length 178;
Best Local Similarity 100.0%; Pred. No. 4.3e-85;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
```

RESULT 2
US-09-770-509-17
; Sequence 17, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
; APPLICANT: Katagiri, F.

; TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES
; FILE REFERENCE: NADII.018A
; CURRENT APPLICATION NUMBER: US/09/770,509
; CURRENT FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Mallomonas splendens
US-09-770-509-17

Query Match 59.5%; Score 527; DB 9; Length 401;
Best Local Similarity 58.4%; Pred. No. 4.6e-47;
Matches 104; Conservative 27; Mismatches 47; Indels 0; Gaps 0;

Qy 1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
Db 100 ARKLSGVEFVCANTDAQHLSTCLTENKLQLGKESTQGLGCGANPESGRRAAEESKEEIAR 159

Qy 61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
Db 160 YIADANMVFITAGMGGGTGTGAAPVVAEVCMEKDILTVAVVTKPFSFEGKHRARLANEGI 219

Qy 121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
Db 220 RSLEDRVDTLIIIPNQNIFKLINASTSMADAFGLADDILLAGVKSITDLMVRPGLINL 277

RESULT 3
US-09-770-509-14
; Sequence 14, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
; APPLICANT: Katagiri, F.
; TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES
; FILE REFERENCE: NADII.018A
; CURRENT APPLICATION NUMBER: US/09/770,509
; CURRENT FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Rickettsia prowazekii
US-09-770-509-14

Query Match 58.5%; Score 518; DB 9; Length 452;
Best Local Similarity 55.9%; Pred. No. 4.8e-46;
Matches 99; Conservative 38; Mismatches 40; Indels 0; Gaps 0;

Qy 2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
Db 36 ANLQGANFVVANTDAQSLEHSLCINKIQLGVSTTRGLGAGASPEVGALAAQESENEIRSS 95

Qy 62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
Db 96 LENSNMVFITAGMGGGTGTGSAPIIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGLI 155

Qy 122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
Db 156 ELQQFVDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINL 212

RESULT 4
US-09-770-509-15
; Sequence 15, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
; APPLICANT: Katagiri, F.
; TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES
; FILE REFERENCE: NADII.018A
; CURRENT APPLICATION NUMBER: US/09/770,509
; CURRENT FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-09-770-509-15

Query Match 57.1%; Score 505; DB 9; Length 508;
Best Local Similarity 58.3%; Pred. No. 1.3e-44;
Matches 102; Conservative 28; Mismatches 45; Indels 0; Gaps 0;

Qy 4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
Db 38 LEGVEFVVANTDAQQLQFAKTDRRIQLGVQITQGLGAGAHPEVGMSAAEESFPEIGEHLD 97

Qy 64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
Db 98 GAHMVFITAGMGGGTGTGAAPIIAKCARERGILTVGVVTKPFHFEGRHRMRLADSGIQEL 157

Qy 124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
Db 158 QRYVDTLIVIPNQNLFRVANERTTFAEAFGMADQVLHSGVRSITDLMVLPGLINL 212

RESULT 5
US-09-770-509-13
; Sequence 13, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
; APPLICANT: Katagiri, F.
; TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES
; FILE REFERENCE: NADII.018A
; CURRENT APPLICATION NUMBER: US/09/770,509
; CURRENT FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Bartonella clarridgeiae
US-09-770-509-13

Query Match 56.0%; Score 496; DB 9; Length 581;
Best Local Similarity 57.7%; Pred. No. 1.4e-43;
Matches 101; Conservative 29; Mismatches 45; Indels 0; Gaps 0;

Qy 4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
Db 38 LQGVDFVVANTDAQALAMSKAERVIQLGAAVTEGLGAGALPEVGRAAADECIDEIIDHLA 97

Qy 64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
Db 98 DSHMVFITAGMGGGTGTGAAPVVANAAREKGILTVGVVTKPFQFEGARRMKTAEAGIEEL 157

Qy 124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
Db 158 QKSVDTLIVIPNQNLFRIANEKTTFSDAFAMADQVLYSGVASITDLMIKEGLINL 212

RESULT 6
US-09-815-242-10639
; Sequence 10639, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.

```
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10639
;   LENGTH: 410
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-815-242-10639

  Query Match             55.9%;  Score 495;  DB 10;  Length 410;
  Best Local Similarity   57.1%;  Pred. No. 1.1e-43;
  Matches  100;  Conservative   28;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  |||| |||  | |   | ||   |:|||||:||:|:::||: :  |   ||
Db         36 VKGVEFITANTDVQALKHSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQVISESLQ 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:||| |||||| |||||||| :|:||| ||||||: ||  ||| | | |  |:  |
Db         96 GADMIFITAGMGGGTGTGAAPVVAKIAKELGALTVGVVTRPFSFEGPKRGRFAAEGIALL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        156 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 210


RESULT 7
US-09-815-242-4909
; Sequence 4909, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4909
;   LENGTH: 411
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(411)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-4909

  Query Match             55.9%;  Score 495;  DB 10;  Length 411;
  Best Local Similarity   57.1%;  Pred. No. 1.1e-43;
  Matches  100;  Conservative   28;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  |||| |||  | |   | ||   |:|||||:||:|:::||: :  |   ||
Db         36 VKGVEFITANTDVQALKHSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQVISESLQ 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:||| |||||| |||||||| :|:||| ||||||: ||  ||| | | |  |:  |
Db         96 GADMIFITAGMGGGTGTGAAPVVAKIAKELGALTVGVVTRPFSFEGPKRGRFAAEGIALL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        156 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 210


RESULT 8
US-09-770-509-16
; Sequence 16, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
;  APPLICANT: Katagiri, F.
;  TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
;  TITLE OF INVENTION:  ANTIMICROBIAL-SPECIFIC BIOCIDES
;  FILE REFERENCE: NADII.018A
;  CURRENT APPLICATION NUMBER: US/09/770,509
;  CURRENT FILING DATE:  2001-01-25
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 407
;   TYPE: PRT
;   ORGANISM: Cyanidioschyzon merolae
US-09-770-509-16

  Query Match             55.4%;  Score 490;  DB 9;  Length 407;
  Best Local Similarity   53.4%;  Pred. No. 3.6e-43;
  Matches   95;  Conservative   34;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              || | ||||:||||| |||  || |::| ||  :|:||||||::|::|: :|||:    ::|
Db        117 ASSLPGVEFLVANTDAQALKMSLCPNRIQLGASLTEGLGAGARPDIGRAAAEEAYETLKR 176

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
                :  ::|||:| |||||| |||||::|  | ||| ||| ||: ||  ||  | : |  |:
Db        177 EFRGVHLLFVTAGMGGGTGTGAAPIIARAAAELGCLTVAVVTKPFHFEGMIRMKTAEQGI 236

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || ::|||::|:||||| :|   |: |:||| || ||  ||: :|||:  ||||||
Db        237 VELTEHVDTMLVIPNQNLFKVASPRTSFLDAFRLADHVLYSGVRSITDLMTVPGLINL 294
```

```
RESULT 9
US-09-712-363-221
; Sequence 221, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
;  APPLICANT: Eisenberg, David
;  APPLICANT:  Rotstein, Sergio H.
;  APPLICANT:  Marcotte, Edward M.
;  TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
;  TITLE OF INVENTION:  INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
;  FILE REFERENCE: 07419-032001
;  CURRENT APPLICATION NUMBER: US/09/712,363
;  CURRENT FILING DATE:  2000-11-13
;  PRIOR APPLICATION NUMBER: PCT/US00/02246
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: 60/179,531
;  PRIOR FILING DATE: 2000-02-01
;  PRIOR APPLICATION NUMBER: 60/117,844
;  PRIOR FILING DATE: 1999-01-29
;  PRIOR APPLICATION NUMBER: 60/118,206,
;  PRIOR FILING DATE: 1999-02-01
;  PRIOR APPLICATION NUMBER: 60/126,593
;  PRIOR FILING DATE: 1999-03-26
;  PRIOR APPLICATION NUMBER: 60/134,093
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/134,092
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/165,124
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/165,086
;  PRIOR FILING DATE: 1999-11-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 221
;   LENGTH: 379
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-09-712-363-221

  Query Match           54.8%;  Score 485;  DB 9;  Length 379;
  Best Local Similarity   54.3%;  Pred. No. 1.1e-42;
  Matches   95;  Conservative   35;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||||  ||| |||  | |  |: :|:| |:|||||: ||:|::::||  | :|: :|:
Db         32 LKGVEFIAINTDAQALLMSDADVKLDVGRDSTRGLGAGADPEVGRKAAEDAKDEIEELLR 91

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| | |||| || |||||:||:|| |||||: ||  ||  |:  |  |:  |
Db         92 GADMVFVTAGEGGGTGTGGAPVVASIARKLGALTVGVVTRPFSFEGKRRSNQAENGIAAL 151

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                :  |||||:||  || : | : ::::||| ||:||| ||:|:|||| |||||:
Db        152 RESCDTLIVIPNDRLLQMGDAAVSLMDAFRSADEVLLNGVQGITDLITTPGLINV 206


RESULT 10
US-09-770-509-11
; Sequence 11, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
;  APPLICANT: Katagiri, F.
;  TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
;  TITLE OF INVENTION:  ANTIMICROBIAL-SPECIFIC BIOCIDES
;  FILE REFERENCE: NADII.018A
;  CURRENT APPLICATION NUMBER: US/09/770,509
;  CURRENT FILING DATE:  2001-01-25
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 583
;   TYPE: PRT
;   ORGANISM: Agrobacterium tumefaciens
US-09-770-509-11

  Query Match           54.5%;  Score 482;  DB 9;  Length 583;
  Best Local Similarity   57.1%;  Pred. No. 4.1e-42;
  Matches  100;  Conservative   26;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:||:|:||||| ||| : |   | || ::|:||||||:||:|: :||:   :|   |
Db         38 LQGVDFVVANTDAQALTMTKADRVIQLGVNVTEGLGAGSQPEVGRAAAEECIDEIIDHLN 97

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::| |:| |||||| |||||||  ||  ||||||||: ||  ||  | |||  |::||
Db         98 GTHMCFVTAGMGGGTGTGAAPVVAQAARNKGILTVGVVTKPFHFEGGRRMRLAEQGIEEL 157

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                | ||||||:||||| :|:  ||  :||  || ||  ||  :|||:|: |||||
Db        158 QKSVDTLIVIPNQNLFRIANDKTTFADAFAMADQVLYSGVACITDLMVKEGLINL 212


RESULT 11
US-09-770-509-10
; Sequence 10, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
;  APPLICANT: Katagiri, F.
;  TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
;  TITLE OF INVENTION:  ANTIMICROBIAL-SPECIFIC BIOCIDES
;  FILE REFERENCE: NADII.018A
;  CURRENT APPLICATION NUMBER: US/09/770,509
;  CURRENT FILING DATE:  2001-01-25
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 420
;   TYPE: PRT
;   ORGANISM: Phytophthora infestans
US-09-770-509-10

  Query Match           54.4%;  Score 481;  DB 9;  Length 420;
  Best Local Similarity   51.7%;  Pred. No. 3.3e-42;
  Matches   92;  Conservative   39;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              |  |:|||||:| ||| | |  :|  ::: :  ::| ||| |: ||:|: :||    :|
Db        129 ARGLQGVEFLVCNTDAQHLRTTLTENRVQMAPELTGGLGCGANPEVGREAAEAAIDEILE 188

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
                :| :||:|:| |||||| ||||||:|  | : ||||| ||: ||| || || :||  |:
Db        189 RVQGANMMFVTAGMGGGTGTGAAPVIAQAALDAGILTVAVVTKPFRFEGNNRAKLAAQGL 248

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                ||   |||::|:|||||  :::: |::::||| ||:|||:||| ::||:| ||||||
Db        249 AELKDSVDTMLVIPNQNLFNMSNERTSLMDAFRMADNVLLDGVKNISDLMVMPGLINL 306


RESULT 12
US-09-815-242-5770
; Sequence 5770, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
```

```
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5770
;   LENGTH: 390
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-5770

  Query Match             54.2%;  Score 480;  DB 10;  Length 390;
  Best Local Similarity   54.3%;  Pred. No. 3.8e-42;
  Matches   95;  Conservative   33;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              :  ||||  ||| |||  | |  || :|: :|:||||||: ||:||::||: :  |:  :|
Db         35 MNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQ 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ::|:|:| |||||| |||||||| :|:|:| ||||||: ||  ||  |   | |||: :
Db         95 GADMVFVTSGMGGGTGTGAAPVVAKIAKEMGALTVGVVTRPFSFEGRKRQTQAAAGVEAM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                ||||||:||  || : |||| |:|||: ||:|| :||:|::|||   | :||
Db        155 KAAVDTLIVIPNDRLLDIVDKSTPMMEAFKEADNVLRQGVQGISDLIAVSGEVNL 209


RESULT 13
US-09-815-242-12723
; Sequence 12723, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12723
;   LENGTH: 392
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-12723

  Query Match             54.2%;  Score 480;  DB 10;  Length 392;
  Best Local Similarity   54.3%;  Pred. No. 3.8e-42;
  Matches   95;  Conservative   33;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              :  ||||  ||| |||  | |  || :|: :|:||||||: ||:||::||: :  |:  :|
Db         35 MNNVEFIAINTDGQALNLSKAESKIQIGEKLTRGLGAGANPEIGKKAAEESREQIEDAIQ 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ::|:|:| |||||| |||||||| :|:|:| ||||||: ||  ||  |   | |||: :
Db         95 GADMVFVTSGMGGGTGTGAAPVVAKIAKEMGALTVGVVTRPFSFEGRKRQTQAAAGVEAM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                ||||||:||  || : |||| |:|||: ||:|| :||:|::|||   | :||
Db        155 KAAVDTLIVIPNDRLLDIVDKSTPMMEAFKEADNVLRQGVQGISDLIAVSGEVNL 209


RESULT 14
US-09-770-509-12
; Sequence 12, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
;  APPLICANT: Katagiri, F.
;  TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
;  TITLE OF INVENTION:  ANTIMICROBIAL-SPECIFIC BIOCIDES
;  FILE REFERENCE: NADII.018A
;  CURRENT APPLICATION NUMBER: US/09/770,509
;  CURRENT FILING DATE:  2001-01-25
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 590
;   TYPE: PRT
;   ORGANISM: Sinorhizobium meliloti
US-09-770-509-12

  Query Match             53.7%;  Score 475;  DB 9;  Length 590;
  Best Local Similarity   55.4%;  Pred. No. 2.2e-41;
  Matches   97;  Conservative   29;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:||:|:|||| |||  : |   | :|  :|:||||||:||:|: :||:   :|   ||
Db         38 LQGVDFVVANTDAQALTMTKAERIIQMGVAVTEGLGAGSQPEVGRAAAEECIDEIIDHLQ 97

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ::| |:| |||||| |||||:||  ||  |||||||: ||  ||  | |:|: |: :|
Db         98 GTHMCFVTAGMGGGTGTGAAPIVAQAARNKGILTVGVVTKPFHFEGGRRMRIADQGISDL 157

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              | ||||||:|||||  :|:  ||  :||  || || ||  :|||:|: |||||
Db        158 QKSVDTLIVIPNQNLFRIANDKTTFADAFAMADQVLYSGVACITDLMVKEGLINL 212


RESULT 15
US-09-754-608-2
; Sequence 2, Application US/09754608
; Patent No. US20020004580A1
; GENERAL INFORMATION:
;  APPLICANT: Fueyo, Joanna Lynn
;  APPLICANT:  Lonetto, Michael A.
;  APPLICANT:  Pearce, Kenneth
;  TITLE OF INVENTION: ftsZ
;  FILE REFERENCE: GM10068
;  CURRENT APPLICATION NUMBER: US/09/754,608
;  CURRENT FILING DATE:  2001-01-04
;  PRIOR APPLICATION NUMBER: 09/120,426
;  PRIOR APPLICATION NUMBER: 1998-07-22
```

```
;  PRIOR APPLICATION NUMBER: 60/055,720
;  PRIOR FILING DATE: 1997-08-12
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 419
;   TYPE: PRT
;   ORGANISM: Streptococcus pneumoniae
US-09-754-608-2

  Query Match             53.3%;  Score 472;  DB 10;  Length 419;
  Best Local Similarity   52.6%;  Pred. No. 2.9e-41;
  Matches   91;  Conservative   35;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          6 GVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDS 65
              ||||| |||| |||  : |   | ||  :|:||||| :||:|:::||: :  :   :  :
Db         38 GVEFIAANTDVQALSSTKAETVIQLGPKLTRGLGAGGQPEVGRKAAEESEETLTEAISGA 97

Qy         66 NMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAK 125
              :|:||| |||||: ||||||:| :|::|| |||||: ||  ||  |: | |: :| :
Db         98 DMVFITAGMGGGSGTGAAPVIARIAKDLGALTVGVVTRPFGFEGSKRGQFAVEGINQLRE 157

Qy        126 YVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :||||::: | ||| : || | :|||   ||:|| :||:|:|||  ||||||
Db        158 HVDTLLIISNNNLLEIVDKKTPLLEALSEADNVLRQGVQGITDLITNPGLINL 210


Search completed: June  2, 2003, 08:52:46
Job time : 236 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2003, 08:33:30 ; Search time 271 Seconds
(without alignments)
423.478 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	885	100.0	178	21	US-09-770-509-2	Sequence 2, Appli
2	527	59.5	336	21	US-09-791-537-32478	Sequence 32478, A
3	527	59.5	401	21	US-09-770-509-17	Sequence 17, Appl
4	527	59.5	401	21	US-09-791-537-130384	Sequence 130384,
5	520	58.8	370	21	US-09-791-537-1723	Sequence 1723, Ap
6	518	58.5	452	21	US-09-770-509-14	Sequence 14, Appl
7	518	58.5	452	21	US-09-791-537-69060	Sequence 69060, A
8	517	58.4	289	21	US-09-791-537-40309	Sequence 40309, A
9	515	58.2	289	21	US-09-791-537-40338	Sequence 40338, A
10	514	58.1	289	21	US-09-791-537-40306	Sequence 40306, A
11	514	58.1	289	21	US-09-791-537-40308	Sequence 40308, A
12	514	58.1	289	21	US-09-791-537-40312	Sequence 40312, A
13	510	57.6	398	21	US-09-791-537-29481	Sequence 29481, A
14	506	57.2	399	21	US-09-791-537-105171	Sequence 105171,
15	506	57.2	399	21	US-09-791-537-132641	Sequence 132641,
16	506	57.2	404	21	US-09-791-537-65811	Sequence 65811, A
17	506	57.2	404	21	US-09-791-537-132535	Sequence 132535,
18	505	57.1	508	21	US-09-770-509-15	Sequence 15, Appl
19	505	57.1	508	21	US-09-791-537-74160	Sequence 74160, A
20	499	56.4	590	21	US-09-791-537-22908	Sequence 22908, A
21	498	56.3	382	21	US-09-791-537-10179	Sequence 10179, A
22	496	56.0	581	21	US-09-770-509-13	Sequence 13, Appl
23	496	56.0	581	21	US-09-791-537-125714	Sequence 125714,
24	495	55.9	410	1	PCT-US02-03987-10639	Sequence 10639, A
25	495	55.9	410	22	US-09-815-242-10639	Sequence 10639, A
26	495	55.9	410	24	US-10-072-851-10639	Sequence 10639, A
27	495	55.9	411	1	PCT-US02-03987-4909	Sequence 4909, Ap
28	495	55.9	411	22	US-09-815-242-4909	Sequence 4909, Ap
29	495	55.9	411	24	US-10-072-851-4909	Sequence 4909, Ap
30	495	55.9	411	27	US-60-269-308-4678	Sequence 4678, Ap
31	495	55.9	435	15	US-09-134-000-4658	Sequence 4658, Ap
32	494	55.8	414	15	US-09-107-532-7295	Sequence 7295, Ap
33	494	55.8	414	15	US-09-107-532A-7295	Sequence 7295, Ap
34	494	55.8	592	21	US-09-791-537-132634	Sequence 132634,
35	493	55.7	372	21	US-09-791-537-1738	Sequence 1738, Ap
36	492	55.6	413	21	US-09-791-537-34221	Sequence 34221, A
37	491	55.5	351	21	US-09-791-537-11835	Sequence 11835, A
38	490	55.4	407	21	US-09-770-509-16	Sequence 16, Appl
39	490	55.4	407	21	US-09-791-537-114519	Sequence 114519,
40	489	55.3	412	21	US-09-791-537-77265	Sequence 77265, A
41	488	55.1	419	21	US-09-791-537-82315	Sequence 82315, A
42	486.5	55.0	318	21	US-09-791-537-31564	Sequence 31564, A
43	486	54.9	290	21	US-09-791-537-1722	Sequence 1722, Ap
44	486	54.9	399	21	US-09-791-537-29477	Sequence 29477, A
45	486	54.9	402	21	US-09-791-537-132563	Sequence 132563,

ALIGNMENTS

RESULT 1
US-09-770-509-2
; Sequence 2, Application US/09770509
; GENERAL INFORMATION:
; APPLICANT: Katagiri, F.
; TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES
; FILE REFERENCE: NADII.018A
; CURRENT APPLICATION NUMBER: US/09/770,509
; CURRENT FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Phytophthora infestans
US-09-770-509-2

Query Match 100.0%; Score 885; DB 21; Length 178;
Best Local Similarity 100.0%; Pred. No. 6.4e-91;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
```

```
Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178


RESULT 2
US-09-791-537-32478
; Sequence 32478, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32478
;   LENGTH: 336
;   TYPE: PRT
;   ORGANISM: Zymomonas mobilis
US-09-791-537-32478

  Query Match             59.5%;  Score 527;  DB 21;  Length 336;
  Best Local Similarity   59.6%;  Pred. No. 4.4e-50;
  Matches  106;  Conservative   29;  Mismatches   43;  Indels    0;  Gaps    0;

Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              || ::||:|||||||| |||  | |  :| ||   |:||||||:||:|| :||:    ||
Db         38 ASGVQGVDFIVANTDAQALNISPAEQRIQLGPTTTQGLGAGSRPEVGKAAAEETIEQIQE 97

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
               |: : | ||  |||||| ||||||:| |||: |||||||: ||  ||  | | | :|:
Db         98 ALEGARMCFIAAGMGGGTGTGAAPVIAKVARDRGILTVGVVTKPFNFEGKRRARSAESGI 157

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :|| |:||||||:|||||  :|: :||  :||: ||:|| :||:|:|||:| ||||||
Db        158 EELQKHVDTLIVIPNQNLFLIANPNTTFKQAFQMADEVLQQGVRGITDLMVCPGLINL 215


RESULT 3
US-09-770-509-17
; Sequence 17, Application US/09770509
; GENERAL INFORMATION:
;  APPLICANT: Katagiri, F.
;  TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
;  TITLE OF INVENTION:  ANTIMICROBIAL-SPECIFIC BIOCIDES
;  FILE REFERENCE: NADII.018A
;  CURRENT APPLICATION NUMBER: US/09/770,509
;  CURRENT FILING DATE:  2001-01-25
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 401
;   TYPE: PRT
;   ORGANISM: Mallomonas splendens
US-09-770-509-17

  Query Match             59.5%;  Score 527;  DB 21;  Length 401;
  Best Local Similarity   58.4%;  Pred. No. 5.8e-50;
  Matches  104;  Conservative   27;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              | :| ||||: |||| | |   |  :|: |||: |:||| |: || |:|:||: | :| |
Db        100 ARKLSGVEFVCANTDAQHLSTCLTENKLQLGKESTQGLGCGANPESGRRAAEESKEEIAR 159

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
               : |:||:||| |||||| |||||||| |  |  |||| ||: ||  || :| |||| |:
Db        160 YIADANMVFITAGMGGGTGTGAAPVVAEVCMEKDILTVAVVTKPFSFEGKHRARLANEGI 219

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               : |  |||||::||||:  | : ||:| :||  |||:|| ||| :|||:||||||||
Db        220 RSLEDRVDTLIIIPNQNIFKLINASTSMADAFGLADDILLAGVKSITDLMVRPGLINL 277


RESULT 4
US-09-791-537-130384
; Sequence 130384, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 130384
;   LENGTH: 401
;   TYPE: PRT
;   ORGANISM: Mallomonas splendens
US-09-791-537-130384

  Query Match             59.5%;  Score 527;  DB 21;  Length 401;
  Best Local Similarity   58.4%;  Pred. No. 5.8e-50;
  Matches  104;  Conservative   27;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              | :| ||||: |||| | |   |  :|: |||: |:||| |: || |:|:||: | :| |
Db        100 ARKLSGVEFVCANTDAQHLSTCLTENKLQLGKESTQGLGCGANPESGRRAAEESKEEIAR 159

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
               : |:||:||| |||||| |||||||| |  |  |||| ||: ||  || :| |||| |:
Db        160 YIADANMVFITAGMGGGTGTGAAPVVAEVCMEKDILTVAVVTKPFSFEGKHRARLANEGI 219

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               : |  |||||::||||:  | : ||:| :||  |||:|| ||| :|||:||||||||
Db        220 RSLEDRVDTLIIIPNQNIFKLINASTSMADAFGLADDILLAGVKSITDLMVRPGLINL 277


RESULT 5
US-09-791-537-1723
; Sequence 1723, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1723
;   LENGTH: 370
;   TYPE: PRT
;   ORGANISM: Clostridium lentocellum
US-09-791-537-1723

  Query Match             58.8%;  Score 520;  DB 21;  Length 370;
  Best Local Similarity   58.9%;  Pred. No. 3.2e-49;
  Matches  103;  Conservative   30;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |||||||  ||| ||| || || || :|: :|:|||||: ||:| :|||: : :|   ::
Db         35 LEGVEFITVNTDHQALARSGAPAKIQIGEKMTRGLGAGANPEIGTKSAEESREEILTAIK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|||||  |||||| ||||||:||:|:| ||||||||: ||  ||  |   |  |: ||
Db         95 GADMLFITAGMGGGTGTGAAPVIASIAKEEGILTVGVVTKPFSFEGRKRMINAEKGIAEL 154
```

```
Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||:|:||  :| : || |||::||  ||||| :||:|:|||| ||:|||
Db        155 KQNVDTLVVIPNDKILQVIDKKTTMVDAFSKADDVLQQGVQGITDLISNPGIINL 209


RESULT 6
US-09-770-509-14
; Sequence 14, Application US/09770509
; GENERAL INFORMATION:
;  APPLICANT: Katagiri, F.
;  TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
;  TITLE OF INVENTION:  ANTIMICROBIAL-SPECIFIC BIOCIDES
;  FILE REFERENCE: NADII.018A
;  CURRENT APPLICATION NUMBER: US/09/770,509
;  CURRENT FILING DATE:  2001-01-25
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 452
;   TYPE: PRT
;   ORGANISM: Rickettsia prowazekii
US-09-770-509-14

  Query Match             58.5%;  Score 518;  DB 21;  Length 452;
  Best Local Similarity   55.9%;  Pred. No. 7.2e-49;
  Matches   99;  Conservative   38;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              : |:|  |:|||||| |:|  ||  :|| ||   |:|||||: ||:|  :|:: : :|:
Db         36 ANLQGANFVVANTDAQSLEHSLCINKIQLGVSTTRGLGAGASPEVGALAAQESENEIRSS 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
              |::|||:||| |||||| ||:||::| :|:||||||||||: ||  || :| : |: |:
Db         96 LENSNMVFITAGMGGGTGTGSAPIIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGLI 155

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || ::||||||:|||||  :|:: ||  :||: |||||  ||:||||:: ||||||
Db        156 ELQQFVDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINL 212


RESULT 7
US-09-791-537-69060
; Sequence 69060, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 69060
;   LENGTH: 452
;   TYPE: PRT
;   ORGANISM: Rickettsia prowazekii
US-09-791-537-69060

  Query Match             58.5%;  Score 518;  DB 21;  Length 452;
  Best Local Similarity   55.9%;  Pred. No. 7.2e-49;
  Matches   99;  Conservative   38;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              : |:|  |:|||||| |:|  ||  :|| ||   |:|||||: ||:|  :|:: : :|:
Db         36 ANLQGANFVVANTDAQSLEHSLCINKIQLGVSTTRGLGAGASPEVGALAAQESENEIRSS 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
              |::|||:||| |||||| ||:||::| :|:||||||||||: ||  || :| : |: |:
Db         96 LENSNMVFITAGMGGGTGTGSAPIIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGLI 155

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || ::||||||:|||||  :|:: ||  :||: |||||  ||:||||:: ||||||
Db        156 ELQQFVDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINL 212


RESULT 8
US-09-791-537-40309
; Sequence 40309, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40309
;   LENGTH: 289
;   TYPE: PRT
;   ORGANISM: Wolbachia sp
US-09-791-537-40309

  Query Match             58.4%;  Score 517;  DB 21;  Length 289;
  Best Local Similarity   56.6%;  Pred. No. 4.8e-49;
  Matches  107;  Conservative   29;  Mismatches   41;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:||||| ||| :||   || || ::|||||||: |:::|| :||:   :|
Db          2 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDVGKGAAEESIDEIMEH 61

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
              ::||:||||| |||||| ||||||:|  |||              |||||||: ||  ||
Db         62 IKDSHMLFITAGMGGGTGTGAAPVIAKAAREARAAVKDRAPKEKKILTVGVVTKPFGFEG 121

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
              | |:|  |::|| ||||||||:|||||  :|:: ||  :||: ||:||  |::|||||
Db        122 VRRMRIAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFSDAFKLADNVLHIGIRGVTDL 181

Qy        170 IVRPGLINL 178
              :| ||||||
Db        182 MVMPGLINL 190


RESULT 9
US-09-791-537-40338
; Sequence 40338, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40338
;   LENGTH: 289
;   TYPE: PRT
;   ORGANISM: Wolbachia sp
US-09-791-537-40338

  Query Match             58.2%;  Score 515;  DB 21;  Length 289;
  Best Local Similarity   56.6%;  Pred. No. 8e-49;
  Matches  107;  Conservative   28;  Mismatches   42;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| ||||||| ||| :||   || || ::|||||||: |:::|| :||:   :|
Db          2 SNLQGVNFIVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDIGKGAAEESIDEIMEH 61
```

```
Qy          62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
               ::||:||||| |||||| ||||||:|  |||              |||||||: ||  ||
Db          62 IRDSHMLFITAGMGGGTGTGAAPVIAKAAREARAVVKDKGAKEKKILTVGVVTKPFGFEG 121

Qy         110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                 | | |  |::|| ||||||||:|||||  :|:: ||  :||: ||:||  |::||||
Db         122 VRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFADAFQLADNVLHIGIRGVTDL 181

Qy         170 IVRPGLINL 178
               :: ||||||
Db         182 MIMPGLINL 190


RESULT 10
US-09-791-537-40306
; Sequence 40306, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40306
;   LENGTH: 289
;   TYPE: PRT
;   ORGANISM: Wolbachia sp
US-09-791-537-40306

  Query Match             58.1%;  Score 514;  DB 21;  Length 289;
  Best Local Similarity   56.1%;  Pred. No. 1e-48;
  Matches  106;  Conservative   29;  Mismatches   42;  Indels   12;  Gaps    1;

Qy           2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
               | |:|| |:||||| ||| :||   || || ::||||||: |::|| :||:   :|
Db           2 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDIGKGAAEESIDEIMEH 61

Qy          62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
               ::||:||||| |||||| ||||||:|  |||              |||||||: ||  ||
Db          62 IRDSHMLFITAGMGGGTGTGAAPVIAEAAREARAVVKDKGAKEKKILTVGVVTKPFGFEG 121

Qy         110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                 | | |  |::|| ||||||||:|||||  :|:: ||  :||: ||:||  |::||||
Db         122 VRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFADAFQLADNVLHIGIRGVTDL 181

Qy         170 IVRPGLINL 178
               :: ||||||
Db         182 MIMPGLINL 190


RESULT 11
US-09-791-537-40308
; Sequence 40308, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40308
;   LENGTH: 289
;   TYPE: PRT
;   ORGANISM: Wolbachia sp
US-09-791-537-40308

  Query Match             58.1%;  Score 514;  DB 21;  Length 289;
  Best Local Similarity   56.1%;  Pred. No. 1e-48;
  Matches  106;  Conservative   29;  Mismatches   42;  Indels   12;  Gaps    1;

Qy           2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
               | |:|| |:||||| ||| :||   || || ::||||||: |::|| :||:   :|
Db           2 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDIGKGAAEESIDEIMEH 61

Qy          62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
               ::||:||||| |||||| ||||||:|  |||              |||||||: ||  ||
Db          62 IRDSHMLFITAGMGGGTGTGAAPVIAKAAREARAVVKDKGAKEKKILTVGVVTKPFGFEG 121

Qy         110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                 | | |  |::|| ||||||||:|||||  :|:: ||  :||: ||:||  |::||||
Db         122 VRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFADAFQLADNVLHIGIRGVTDL 181

Qy         170 IVRPGLINL 178
               :: ||||||
Db         182 MIMPGLINL 190


RESULT 12
US-09-791-537-40312
; Sequence 40312, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40312
;   LENGTH: 289
;   TYPE: PRT
;   ORGANISM: Wolbachia sp
US-09-791-537-40312

  Query Match             58.1%;  Score 514;  DB 21;  Length 289;
  Best Local Similarity   56.1%;  Pred. No. 1e-48;
  Matches  106;  Conservative   29;  Mismatches   42;  Indels   12;  Gaps    1;

Qy           2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
               | |:|| |:||||| ||| :||   || || ::||||||: |::|| :||:   :|
Db           2 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDIGKGAAEESIDEIMEH 61

Qy          62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
               ::||:||||| |||||| ||||||:|  |||              |||||||: ||  ||
Db          62 IRDSHMLFITAGMGGGTGTGAAPVIAKAAREARAVVKDKGAKEKKILTVGVVTKPFGFEG 121

Qy         110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                 | | |  |::|| ||||||||:|||||  :|:: ||  :||: ||:||  |::||||
Db         122 VRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFADAFQLADNVLHIGIRGVTDL 181

Qy         170 IVRPGLINL 178
               :: ||||||
Db         182 MIMPGLINL 190


RESULT 13
US-09-791-537-29481
; Sequence 29481, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
```

```
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29481
;   LENGTH: 398
;   TYPE: PRT
;   ORGANISM: Wolbachia sp
US-09-791-537-29481

  Query Match             57.6%;  Score 510;  DB 21;  Length 398;
  Best Local Similarity   56.1%;  Pred. No. 4.7e-48;
  Matches  106;  Conservative   29;  Mismatches   42;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:|||| ||| :||   || || ::|||||||: |::|| :||:   :|
Db         36 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDVGKGAAEESIDEIMEH 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
              ::||:|||| |||||| ||||||:|  |||              |||||||: ||  ||
Db         96 IKDSHMLFITAGMGGGTGTGAAPVIAKAAREARAAVKDRAPKEKKILTVGVVTKPFGFEG 155

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                |  :|  |::|| |||||||:|||||  :|:: ||  :||: ||:||  |::|||||
Db        156 VRRMPIAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFSDAFKLADNVLHIGIRGVTDL 215

Qy        170 IVRPGLINL 178
              :| ||||||
Db        216 MVMPGLINL 224


RESULT 14
US-09-791-537-105171
; Sequence 105171, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 105171
;   LENGTH: 399
;   TYPE: PRT
;   ORGANISM: Borrelia burgdorferi
US-09-791-537-105171

  Query Match             57.2%;  Score 506;  DB 21;  Length 399;
  Best Local Similarity   58.7%;  Pred. No. 1.3e-47;
  Matches  101;  Conservative   27;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          7 VEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDSN 66
              ||||||||| |||  |:|| || ||  :| ||||| |||:|: :||:    |:  |  ::
Db         47 VEFIVANTDLQALQTSIAPIKIALGAKVTAGLGAGGKPEIGQAAAEEDIDVIRNHLSGAD 106

Qy         67 MLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAKY 126
              |:||| |||||| ||||||:| ||:||||||||||: ||: ||| : |||  |:  | |
Db        107 MVFITAGMGGGTGTGAAPVIAQVAKELGILTVGVVTKPFKFEGPKKLRLAEQGINNLRKS 166

Qy        127 VDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              |||||::||| || : || ||: :||: ||||| ||:|:  ||:  | :|:
Db        167 VDTLIIIPNQKLLTVVDKRTTIKDAFKRADDVLRMGVQGIAGLIIEHGEVNI 218


RESULT 15
US-09-791-537-132641
; Sequence 132641, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132641
;   LENGTH: 399
;   TYPE: PRT
;   ORGANISM: Borrelia burgdorferi
US-09-791-537-132641

  Query Match             57.2%;  Score 506;  DB 21;  Length 399;
  Best Local Similarity   58.7%;  Pred. No. 1.3e-47;
  Matches  101;  Conservative   27;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          7 VEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDSN 66
              ||||||||| |||  |:|| || ||  :| ||||| |||:|: :||:    |:  |  ::
Db         47 VEFIVANTDLQALQTSIAPIKIALGAKVTAGLGAGGKPEIGQAAAEEDIDVIRNHLSGAD 106

Qy         67 MLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAKY 126
              |:||| |||||| ||||||:| ||:||||||||||: ||: ||| : |||  |:  | |
Db        107 MVFITAGMGGGTGTGAAPVIAQVAKELGILTVGVVTKPFKFEGPKKLRLAEQGINNLRKS 166

Qy        127 VDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              |||||::||| || : || ||: :||: ||||| ||:|:  ||:  | :|:
Db        167 VDTLIIIPNQKLLTVVDKRTTIKDAFKRADDVLRMGVQGIAGLIIEHGEVNI 218


Search completed: June  2, 2003, 08:47:09
Job time : 276 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2003, 08:41:25 ; Search time 37 Seconds
(without alignments)
1012.472 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1005382 seqs, 210457600 residues

Total number of hits satisfying chosen parameters: 1005382

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	510	57.6	391	6	US-10-282-122A-60624	Sequence 60624, A
2	507	57.3	373	6	US-10-282-122A-51618	Sequence 51618, A
3	506	57.2	404	6	US-10-282-122A-47081	Sequence 47081, A
4	505	57.1	386	6	US-10-282-122A-45663	Sequence 45663, A
5	505	57.1	386	6	US-10-282-122A-46846	Sequence 46846, A
6	505	57.1	394	6	US-10-282-122A-53331	Sequence 53331, A
7	495	55.9	410	6	US-10-282-122A-42469	Sequence 42469, A
8	495	55.9	435	5	US-09-134-000C-4658	Sequence 4658, Ap
9	495	55.9	435	5	US-09-134-000C-4658	Sequence 4658, Ap
10	494	55.8	413	6	US-10-282-122A-57592	Sequence 57592, A
11	494	55.8	414	6	US-10-417-884-7295	Sequence 7295, Ap
12	492	55.6	387	6	US-10-282-122A-61968	Sequence 61968, A
13	490	55.4	369	6	US-10-282-122A-52722	Sequence 52722, A
14	490	55.4	410	1	PCT-US02-36122-84	Sequence 84, Appl
15	490	55.4	410	1	PCT-US02-36123-5202	Sequence 5202, Ap
16	486	54.9	396	6	US-10-156-761-13651	Sequence 13651, A
17	486	54.9	411	6	US-10-282-122A-54015	Sequence 54015, A
18	485	54.8	379	6	US-10-282-122A-62334	Sequence 62334, A
19	485	54.8	379	6	US-10-282-122A-64648	Sequence 64648, A
20	482	54.5	393	6	US-10-282-122A-71506	Sequence 71506, A
21	481	54.4	379	6	US-10-282-122A-63791	Sequence 63791, A
22	480	54.2	390	5	US-09-950-084-5515	Sequence 5515, Ap
23	480	54.2	390	6	US-10-282-122A-44008	Sequence 44008, A
24	480	54.2	394	6	US-10-092-411A-4834	Sequence 4834, Ap
25	480	54.2	394	6	US-10-282-122A-70732	Sequence 70732, A
26	472	53.3	419	6	US-10-282-122A-74113	Sequence 74113, A
27	471	53.2	398	6	US-10-282-122A-67907	Sequence 67907, A
28	468	52.9	395	6	US-10-282-122A-69702	Sequence 69702, A
29	466	52.7	421	1	PCT-US02-32727-21663	Sequence 21663, A
30	466	52.7	421	5	US-09-978-825-21663	Sequence 21663, A
31	466	52.7	421	6	US-10-057-498-21663	Sequence 21663, A
32	465.5	52.6	356	6	US-10-282-122A-63205	Sequence 63205, A
33	465	52.5	418	6	US-10-282-122A-76497	Sequence 76497, A
34	464	52.4	394	6	US-10-282-122A-66647	Sequence 66647, A
35	464	52.4	400	6	US-10-366-683-24128	Sequence 24128, A
36	464	52.4	400	6	US-10-419-128-24128	Sequence 24128, A
37	464	52.4	439	6	US-10-282-122A-74650	Sequence 74650, A
38	460	52.0	383	6	US-10-282-122A-77931	Sequence 77931, A
39	459	51.9	425	5	US-09-500-495A-6	Sequence 6, Appli
40	459	51.9	434	6	US-10-282-122A-72045	Sequence 72045, A
41	455.5	51.5	391	6	US-10-282-122A-44660	Sequence 44660, A
42	453	51.2	383	6	US-10-282-122A-75194	Sequence 75194, A
43	453	51.2	383	6	US-10-282-122A-76027	Sequence 76027, A
44	453	51.2	414	6	US-10-424-599-177675	Sequence 177675,
45	452	51.1	383	6	US-10-282-122A-42973	Sequence 42973, A

ALIGNMENTS

RESULT 1
US-10-282-122A-60624
; Sequence 60624, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60624
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60624

Query Match 57.6%; Score 510; DB 6; Length 391;
Best Local Similarity 56.0%; Pred. No. 2.2e-39;
Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  ||| |||  : |  |: :|  :|:|||||: ||:||::||: :  |:  |:
Db         35 VQGVEFISVNTDAQALNLAKAETKLQIGTKLTRGLGAGAVPEIGKKAAEESREQIEEALK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               |:|:|:| |||||| ||||||:| :|:|:| ||||||: ||  ||| ||: |  | : :
Db         95 GSDMVFVTAGMGGGTGTGAAPVIAQIAKEMGALTVGVVTRPFGFEGPKRTKQALTGTEAM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  || : ||:| |||||| ||:|| :||:|::|||  ||||||
Db        155 KEAVDTLIVIPNDRLLQIVDKNTPMLEAFREADNVLRQGVQGISDLIAVPGLINL 209
```

RESULT 2
US-10-282-122A-51618
; Sequence 51618, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51618
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51618

Query Match 57.3%; Score 507; DB 6; Length 373;
Best Local Similarity 56.0%; Pred. No. 3.9e-39;
Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |: ||||  ||| |||  | |  || :|  :|||||||: ||:|:::||: | :| : ::
Db         35 LKNVEFIAINTDKQALALSQASQKIQIGDKLTKGLGAGANPEIGQKAAEESKDEISQAIK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ::|:||| |||||| |||||||| :|: :|||||||: ||  ||  |   | :|:| |
Db         95 GADMVFITAGMGGGTGTGAAPVVAEIAKSMGILTVGVVTKPFPFEGRKRMLHAESGIKTL 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||: :||: |||: || ||::|||: ||||| :||:|::|||  |||:||
Db        155 KERVDTLVTIPNERLLAIVDKKTTLVEAFKSADDVLRQGVQGISDLITIPGLVNL 209
```

RESULT 3
US-10-282-122A-47081
; Sequence 47081, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47081
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-10-282-122A-47081

Query Match 57.2%; Score 506; DB 6; Length 404;
Best Local Similarity 58.7%; Pred. No. 5.5e-39;
Matches 101; Conservative 27; Mismatches 44; Indels 0; Gaps 0;

```
Qy          7 VEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDSN 66
              ||||||||| |||  |:|| || ||  :| ||||| |||:|: :|||:    |:  |  ::
Db         52 VEFIVANTDLQALQTSIAPIKIALGAKVTAGLGAGGKPEIGQAAAEEDIDVIRNHLSGAD 111

Qy         67 MLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAKY 126
              |:||| |||||| ||||||:| ||:||||||||||: ||: ||| : |||  |:  | |
Db        112 MVFITAGMGGGTGTGAAPVIAQVAKELGILTVGVVTKPFKFEGPKKLRLAEQGINNLRKS 171

Qy        127 VDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              ||||||::||| || : || ||: :|||: |||||  ||:|:  ||: .| :|:
```

```
Db        172 VDTLIIIPNQKLLTVVDKRTTIKDAFKRADDVLRMGVQGIAGLIIEHGEVNI 223


RESULT 4
US-10-282-122A-45663
; Sequence 45663, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45663
;   LENGTH: 386
;   TYPE: PRT
;   ORGANISM: Bacillus anthracis
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (269)..(269)
;   OTHER INFORMATION: X=any amino acid
US-10-282-122A-45663

  Query Match             57.1%;  Score 505;  DB 6;  Length 386;
  Best Local Similarity   57.1%;  Pred. No. 6.4e-39;
  Matches  100;  Conservative   30;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::||:||  ||| |||  | |  |: :|  :|:|||||: ||:||::||: |  ||  |:
Db         35 VQGVDFIAVNTDAQALNLSKAETKMQIGGKLTRGLGAGANPEVGKKAAEESKEQIQEALR 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| |||||| |||||||| ||:||| ||||||: ||  ||  |   | :|:
Db         95 GADMVFVTAGMGGGTGTGAAPVVAQVAKELGALTVGVVTRPFTFEGRKRATQAASGIAAF 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  || : ||:| |||||| ||:|| :||:|::|||  ||||||
Db        155 KENVDTLIVIPNDRLLEIVDKNTPMLEAFREADNVLRQGVQGISDLIATPGLINL 209


RESULT 5
US-10-282-122A-46846
; Sequence 46846, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46846
;   LENGTH: 386
;   TYPE: PRT
;   ORGANISM: Bacillus anthracis
US-10-282-122A-46846

  Query Match             57.1%;  Score 505;  DB 6;  Length 386;
  Best Local Similarity   57.1%;  Pred. No. 6.4e-39;
  Matches  100;  Conservative   30;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::||:||  ||| |||  | |  |: :|  :|:|||||: ||:||::||: |  ||  |:
Db         35 VQGVDFIAVNTDAQALNLSKAETKMQIGGKLTRGLGAGANPEVGKKAAEESKEQIQEALR 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| |||||| |||||||| ||:||| ||||||: ||  ||  |   | :|:
Db         95 GADMVFVTAGMGGGTGTGAAPVVAQVAKELGALTVGVVTRPFTFEGRKRATQAASGIAAF 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  || : ||:| |||||| ||:|| :||:|::|||  ||||||
Db        155 KENVDTLIVIPNDRLLEIVDKNTPMLEAFREADNVLRQGVQGISDLIATPGLINL 209


RESULT 6
US-10-282-122A-53331
; Sequence 53331, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
```

```
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53331
;   LENGTH: 394
;   TYPE: PRT
;   ORGANISM: Clostridium difficile
US-10-282-122A-53331

  Query Match             57.1%;  Score 505;  DB 6;  Length 394;
  Best Local Similarity   55.9%;  Pred. No. 6.5e-39;
  Matches   99;  Conservative   35;  Mismatches   43;  Indels    0;  Gaps    0;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              :||:|||||  ||| |||  | | :|: :|: :|:|||||: ||:|||:||: | :| ::
Db         42 AQLKGVEFISVNTDKQALYTSKAEYKVQIGEKLTRGLGAGANPEVGKRAAEESKDEIVKL 101

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
              || ::|:|:| |||||| |||||||| :|:|:|||||||: ||  ||  | : |  |:
Db        102 LQGADMVFVTAGMGGGTGTGAAPVVAGLAKEMGILTVGVVTKPFAFEGKIRMKNAEGGIA 161

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              ||   ||||| :||  || :  |:|:||:||  ||||| :|:: ::|||   |||||
Db        162 ELKSKVDTLITIPNDRLLQIVQKNTSMLDAFAVADDVLKQGIQSISDLIAVEGLINL 218


RESULT 7
US-10-282-122A-42469
; Sequence 42469, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42469
;   LENGTH: 410
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-10-282-122A-42469

  Query Match             55.9%;  Score 495;  DB 6;  Length 410;
  Best Local Similarity   57.1%;  Pred. No. 6e-38;
  Matches  100;  Conservative   28;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::||||| |||| |||  | |   | ||   |:||||||:||:|:::||: :  |   ||
Db         36 VKGVEFITANTDVQALKHSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQVISESLQ 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ::|:||| |||||| |||||||| :|:||| ||||||: ||  ||| | | |  |:  |
Db         96 GADMIFITAGMGGGTGTGAAPVVAKIAKELGALTVGVVTRPFSFEGPKRGRFAAEGIALL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        156 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 210


RESULT 8
US-09-134-000C-4658
; Sequence 4658, Application US/09134000C
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4658
;   LENGTH: 435
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-4658

  Query Match             55.9%;  Score 495;  DB 5;  Length 435;
```

```
  Best Local Similarity   57.1%;  Pred. No. 6.5e-38;
  Matches  100;  Conservative   28;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  |||| |||  | |   | ||   |:||||||:||:|:::||: :  |   ||
Db         61 VKGVEFITANTDVQALKHSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQVISESLQ 120

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:||| |||||| |||||||| :|:||| ||||||: ||  ||| | | |  |:  |
Db        121 GADMIFITAGMGGGTGTGAAPVVAKIAKELGALTVGVVTRPFSFEGPKRGRFAAEGIALL 180

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        181 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 235


RESULT 9
US-09-134-000C-4658
; Sequence 4658, Application US/09134000C
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4658
;   LENGTH: 435
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-4658

  Query Match           55.9%;  Score 495;  DB 5;  Length 435;
  Best Local Similarity   57.1%;  Pred. No. 6.5e-38;
  Matches  100;  Conservative   28;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  |||| |||  | |   | ||   |:||||||:||:|:::||: :  |   ||
Db         61 VKGVEFITANTDVQALKHSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQVISESLQ 120

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:||| |||||| |||||||| :|:||| ||||||: ||  ||| | | |  |:  |
Db        121 GADMIFITAGMGGGTGTGAAPVVAKIAKELGALTVGVVTRPFSFEGPKRGRFAAEGIALL 180

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        181 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 235


RESULT 10
US-10-282-122A-57592
; Sequence 57592, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER:.US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57592
;   LENGTH: 413
;   TYPE: PRT
;   ORGANISM: Enterococcus faecium
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (343)..(343)
;   OTHER INFORMATION: X=any amino acid
US-10-282-122A-57592

  Query Match           55.8%;  Score 494;  DB 6;  Length 413;
  Best Local Similarity   55.4%;  Pred. No. 7.5e-38;
  Matches   97;  Conservative   31;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  |||| |||  | |   | ||   |:||||||:||:|:::||: :  ::  |
Db         36 VKGVEFITANTDVQALKNSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQSLREALD 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:||| |||||| |||||:|| :|:||| ||||||: ||  ||| | | |  |:  |
Db         96 GADMIFITAGMGGGTGTGAAPIVAGIAKELGALTVGVVTRPFTFEGPKRGRFAAEGIARL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        156 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 210


RESULT 11
US-10-417-884-7295
; Sequence 7295, Application US/10417884
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/417,884
;             FILING DATE: 17-Apr-2003
```

```
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;    INFORMATION FOR SEQ ID NO: 7295:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 414 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...414
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7295:
US-10-417-884-7295

  Query Match             55.8%;  Score 494;  DB 6;  Length 414;
  Best Local Similarity   55.4%;  Pred. No. 7.5e-38;
  Matches   97;  Conservative   31;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  |||| |||  | |   | ||   |:||||||:||:|:::||: :  ::  |
Db         37 VKGVEFITANTDVQALKNSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQSLREALD 96

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:||| |||||| |||||:|| :|:||| |||||: ||  ||| | | |  |:  |
Db         97 GADMIFITAGMGGGTGTGAAPIVAGIAKELGALTVGVVTRPFTFEGPKRGRFAAEGIARL 156

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        157 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 211


RESULT 12
US-10-282-122A-61968
; Sequence 61968, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61968
;   LENGTH: 387
;   TYPE: PRT
;   ORGANISM: Mycobacterium avium
US-10-282-122A-61968

  Query Match             55.6%;  Score 492;  DB 6;  Length 387;
  Best Local Similarity   54.9%;  Pred. No. 1.1e-37;
  Matches   96;  Conservative   35;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||||  ||| |||  | |  |: :|:| |:|||||: ||:|:::||  | :|: :|:
Db         33 LKGVEFIAINTDAQALLMSDADVKLDVGRDSTRGLGAGADPEVGRKAAEDAKDEIEELLR 92

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| | |||| || ||||||:||:|| ||||||: ||  ||  |:  | ||:  |
Db         93 GADMVFVTAGEGGGTGTGGAPVVASIARKLGALTVGVVTRPFSFEGKRRSNQAEAGINAL 152

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               :  |||||:||  || : | : ::::||| ||:||| ||:|:|||| |||||:
Db        153 RESCDTLIVIPNDRLLQMGDAAVSLMDAFRSADEVLLNGVQGITDLITTPGLINV 207


RESULT 13
US-10-282-122A-52722
; Sequence 52722, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
```

```
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52722
;   LENGTH: 369
;   TYPE: PRT
;   ORGANISM: Clostridium botulinum
US-10-282-122A-52722

  Query Match             55.4%;  Score 490;  DB 6;  Length 369;
  Best Local Similarity   53.1%;  Pred. No. 1.5e-37;
  Matches   93;  Conservative   38;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |: ||||  ||| |||  | |  || :|  :|||||||: ||:||::||: | :| : ::
Db         35 LKNVEFIAINTDKQALMLSQASQKIQIGDKLTKGLGAGANPEIGKKAAEESKEEISQAIK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:||| |||||| ||||||:| :|: :||||||||: ||  ||  |   |  |:  |
Db         95 GADMVFITAGMGGGTGTGAAPVIAEIAKSMGILTVGVVTKPFPFEGRKRLLHAEMGINTL 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                : ||||: :||: ||:: || |:::::|: ||||| :||:|::|||  |||:||
Db        155 KERVDTLVTIPNERLLSIVDKKTSLMDSFKLADDVLRQGVQGISDLITIPGLVNL 209


RESULT 14
PCT-US02-36122-84
; Sequence 84, Application PC/TUS0236122
; GENERAL INFORMATION:
;  APPLICANT: Murphy, Ellen and Projan, Stephen, j.
;  TITLE OF INVENTION: Alloiococcus otitidis Infectious Disease Targets
;  FILE REFERENCE: Application 1
;  CURRENT APPLICATION NUMBER: PCT/US02/36122
;  CURRENT FILING DATE:  2002-11-25
;  NUMBER OF SEQ ID NOS: 106
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 84
;   LENGTH: 410
;   TYPE: PRT
;   ORGANISM: Alloiococcus otitidis
PCT-US02-36122-84

  Query Match             55.4%;  Score 490;  DB 1;  Length 410;
  Best Local Similarity   52.6%;  Pred. No. 1.8e-37;
  Matches   92;  Conservative   37;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              :|||||||||||| |||  : |  || ||: :|:|||||: ||:|::|||: :  |   |:
Db         36 VEGVEFIVANTDVQALDANRAETKIQLGEKLTRGLGAGANPEVGRKSAEESEETIAEALE 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| |||||| || | ::| :|:| | |||||:: ||  ||| | | |  |: ::
Db         96 GADMVFVTAGMGGGTGTGGAGIIARIAKEQGALTVGVITRPFTFEGPKRGRFAAEGIAQM 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                ::|||||: : |  || : || | |:|||| ||:|| :||:|::|||  || :||
Db        156 REHVDTLVTISNNRLLEIVDKKTPMMEAFREADNVLRQGVQGISDLITNPGYVNL 210


RESULT 15
PCT-US02-36123-5202
; Sequence 5202, Application PC/TUS0236123
; GENERAL INFORMATION:
;  APPLICANT: American Cyanamid Company, and Fletcher, Leah D., McMichael, John C.,
;  APPLICANT:  Russell, David P., and Zagursky, Robert J.
;  TITLE OF INVENTION: Alloiococcus otitidis Open Reading Frames (ORFs) Encoding Polypep
;  TITLE OF INVENTION:  Antigens, Immunogenic Compositions and Uses Thereof
;  FILE REFERENCE: Application 1
;  CURRENT APPLICATION NUMBER: PCT/US02/36123
;  CURRENT FILING DATE:  2003-01-02
;  NUMBER OF SEQ ID NOS: 6653
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5202
;   LENGTH: 410
;   TYPE: PRT
;   ORGANISM: Alloiococcus otitidis
PCT-US02-36123-5202

  Query Match             55.4%;  Score 490;  DB 1;  Length 410;
  Best Local Similarity   52.6%;  Pred. No. 1.8e-37;
  Matches   92;  Conservative   37;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              :|||||||||||| |||  : |  || ||: :|:|||||: ||:|::|||: :  |   |:
Db         36 VEGVEFIVANTDVQALDANRAETKIQLGEKLTRGLGAGANPEVGRKSAEESEETIAEALE 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| |||||| || | ::| :|:| | |||||:: ||  ||| | | |  |: ::
Db         96 GADMVFVTAGMGGGTGTGGAGIIARIAKEQGALTVGVITRPFTFEGPKRGRFAAEGIAQM 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                ::|||||: : |  || : || | |:|||| ||:|| :||:|::|||  || :||
Db        156 REHVDTLVTISNNRLLEIVDKKTPMMEAFREADNVLRQGVQGISDLITNPGYVNL 210


Search completed: June  2, 2003, 08:48:40
Job time : 42 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2003, 07:18:30 ; Search time 40 Seconds
(without alignments)
427.798 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	524	59.2	452	2	G97826	cell division prot
2	518	58.5	452	2	F71672	cell division prot
3	511	57.7	382	2	F83969	cell-division init
4	510	57.6	391	2	AH1328	cell-division init
5	510	57.6	392	2	AH1699	cell-division init
6	510	57.6	398	2	S35264	cell division prot
7	507	57.3	373	2	H97108	cell division GTPa
8	506	57.2	404	2	C70137	cell division prot
9	505	57.1	508	2	C87564	cell division prot
10	498	56.3	382	1	I39848	cell division init
11	493	55.7	417	2	E86858	cell division prot
12	486	54.9	399	1	S60765	cell division prot
13	486	54.9	402	2	JE0282	cell division prot
14	485	54.8	379	2	B70579	probable cell divi
15	482	54.5	583	2	C97610	cell division prot
16	482	54.5	583	2	AG2832	cell division prot
17	481	54.4	373	2	JC7087	ftsZ protein - Ach
18	481	54.4	379	2	G87023	cell division prot
19	480	54.2	390	1	S58814	cell division prot
20	476	53.8	538	2	AC3325	cell division prot
21	475	53.7	590	1	A38119	cell division prot
22	472	53.3	419	2	H95193	cell division prot
23	472	53.3	419	2	E98060	cell division prot
24	471	53.2	400	2	T10476	cell division prot
25	468	52.9	379	1	JC4289	cell division prot
26	468	52.9	428	2	AC2288	cell division prot
27	468	52.9	428	4	S46932	hypothetical cell-
28	467	52.8	413	2	B82760	cell division prot
29	467	52.8	430	2	S77393	cell division prot
30	465	52.5	418	2	H71331	probable cell divi
31	464	52.4	394	2	H83093	cell division prot
32	463	52.3	371	2	E75494	cell division prot
33	463	52.3	397	2	E84778	plastid division p
34	463	52.3	478	2	JC7770	chloroplast divisi
35	462	52.2	464	2	T51090	plastid division p
36	461	52.1	384	2	A84955	cell division prot
37	461	52.1	458	2	T51089	plastid division p
38	460.5	52.0	371	2	T51695	cell division prot
39	460	52.0	383	2	AE0069	cell division prot
40	459	51.9	345	1	S58854	cell division prot
41	458	51.8	473	2	T49028	plastid division p
42	455	51.4	399	2	G90102	cell division prot
43	453	51.2	383	2	AB0519	cell division prot
44	452	51.1	383	1	CEECZ	cell division prot
45	452	51.1	383	2	C90641	cell division prot

ALIGNMENTS

RESULT 1
G97826
cell division protein ftsZ [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 17-May-2002
C;Accession: G97826
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: G97826
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-452 <KUR>
A;Cross-references: GB:AE006914; PIDN:AAL03553.1; PID:g15620131; GSPDB:GN00173
C;Genetics:
A;Gene: ftsZ
C;Superfamily: cell division protein ftsZ

Query Match 59.2%; Score 524; DB 2; Length 452;
Best Local Similarity 56.2%; Pred. No. 2.2e-38;
Matches 100; Conservative 38; Mismatches 40; Indels 0; Gaps 0;

```
Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              :: |:|  |:|||||| |:|  ||  :|| ||   |:|||||: ||:|  :|:: : :|:
Db         35 SANLQGANFVVANTDAQSLEHSLCTNKIQLGVSTTRGLGAGASPEVGALAAQESESEIRN 94

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
               |::|||:||| |||||| ||:|||:| :|:|||||||||: ||  || :| : |: |:
Db         95 YLENSNMVFITAGMGGGTGTGSAPVIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGL 154

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               || ::|||||:|||||  :|:: ||  :||: |||||  ||:||||:: ||||||
Db        155 IELQQFVDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINL 212
```

RESULT 2
F71672
cell division protein ftsz (ftsZ) RP666 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C;Accession: F71672
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: F71672
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-452 <AND>
A;Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAA15104.1; PID:e134
A;Experimental source: strain Madrid E

C;Genetics:
A;Gene: ftsZ; RP666
C;Superfamily: cell division protein ftsZ

Query Match 58.5%; Score 518; DB 2; Length 452;
Best Local Similarity 55.9%; Pred. No. 7.2e-38;
Matches 99; Conservative 38; Mismatches 40; Indels 0; Gaps 0;

```
Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              : |:|  |:|||||| |:|  ||  :|| ||   |:|||||: ||:|  :|:: : :|:
Db         36 ANLQGANFVVANTDAQSLEHSLCINKIQLGVSTTRGLGAGASPEVGALAAQESENEIRSS 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
              |::|||:||| |||||| ||:||::| :|:|||||||||: ||  || :| : |: |:
Db         96 LENSNMVFITAGMGGGTGTGSAPIIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGLI 155

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || ::|||||:|||||  :|:: ||  :||: |||||  ||:||||:: ||||||
Db        156 ELQQFVDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINL 212
```

RESULT 3
F83969
cell-division initiation protein (septum formation) ftsZ [imported] - Bacillus halodurans
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: F83969
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: F83969
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-382 <STO>
A;Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BAB06277.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: ftsZ
C;Superfamily: cell division protein ftsZ

Query Match 57.7%; Score 511; DB 2; Length 382;
Best Local Similarity 58.3%; Pred. No. 2.5e-37;
Matches 102; Conservative 30; Mismatches 43; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:||:||  ||| |||  | |  |: ||  :|:|||||: ||:||::||: :  |:  ||
Db         35 LQGVDFISVNTDAQALHLSKAEVKLQLGGKLTRGLGAGANPEIGKKAAEESREQIEEALQ 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:||| |||||| ||||||:| ||:|:| ||||||: ||  ||  |:  | ||:  |
Db         95 GADMVFITAGMGGGTGTGAAPVIAEVAKEIGALTVGVVTRPFTFEGRKRSTQAAAGIAAL 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  || : ||:| |||||| ||:|| :|||:|::||| ||||||
Db        155 KEKVDTLIVIPNDRLLEIVDKNTPMLEAFREADNVLRQGVQGISDLIATPGLINL 209
```

RESULT 4
AH1328
cell-division initiation protein FtsZ homolog ftsZ [imported] - Listeria monocytogenes (
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AH1328
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1328
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-391 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00110.1; PID:g16411502; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: ftsZ
C;Superfamily: cell division protein ftsZ

Query Match 57.6%; Score 510; DB 2; Length 391;
Best Local Similarity 56.0%; Pred. No. 3.1e-37;
Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  ||| |||  : |  |: :|  :|:|||||: ||:||::||: :  |:  |:
Db         35 VQGVEFISVNTDAQALNLAKAETKLQIGTKLTRGLGAGAVPEIGKKAAEESREQIEEALK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               |:|:|:| |||||| ||||||:| :|:|:| ||||||: ||  ||| ||: |  | : :
Db         95 GSDMVFVTAGMGGGTGTGAAPVIAQIAKEMGALTVGVVTRPFGFEGPKRTKQALTGTEAM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  || : ||:| |||||| ||:|| :|||:|::||| ||||||
Db        155 KEAVDTLIVIPNDRLLQIVDKNTPMLEAFREADNVLRQGVQGISDLIAVPGLINL 209
```

RESULT 5
AH1699
cell-division initiation protein FtsZ homolog ftsZ [imported] - Listeria innocua (str
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AH1699
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1699
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-392 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97368.1; PID:g16414652; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: ftsZ
C;Superfamily: cell division protein ftsZ

Query Match 57.6%; Score 510; DB 2; Length 392;
Best Local Similarity 56.0%; Pred. No. 3.1e-37;
Matches 98; Conservative 35; Mismatches 42; Indels 0; Gaps 0;

```
Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  ||| |||  : |  |: :|  :|:|||||: ||:||::||: :  |:  |:
Db         35 VQGVEFISVNTDAQALNLAKAETKLQIGTKLTRGLGAGAVPEIGKKAAEESREQIEEALK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               |:|:|:| |||||| ||||||:| :|:|:| ||||||: ||  ||| ||: |  | : :
Db         95 GSDMVFVTAGMGGGTGTGAAPVIAQIAKEMGALTVGVVTRPFGFEGPKRTKQALTGTEAM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  || : ||:| |||||| ||:|| :|||:|::||| ||||||
Db        155 KEAVDTLIVIPNDRLLQIVDKNTPMLEAFREADNVLRQGVQGISDLIAVPGLINL 209
```

RESULT 6
S35264
cell division protein ftsZ homolog - Wolbachia sp. (fragment)
C;Species: Wolbachia sp.
C;Date: 09-Dec-1993 #sequence_revision 12-Apr-1996 #text_change 26-Aug-1999
C;Accession: S35264; S35404

```
R;Holden, P.R.; Brookfield, J.F.Y.; Jones, P.
Mol. Gen. Genet. 240, 213-220, 1993
A;Title: Cloning and characterization of an ftsZ homologue from a bacterial symbiont of
A;Reference number: S35264; MUID:93360900; PMID:7689140
A;Accession: S35264
A;Molecule type: DNA
A;Residues: 1-398 <HOL>
A;Cross-references: EMBL:X71906; NID:g311274; PIDN:CAA50724.1; PID:g311275
A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in lacking 284-Gl
C;Genetics:
A;Gene: ftsZ
A;Cross-references: FlyBase:FBgn0013356
C;Superfamily: cell division protein ftsZ
C;Keywords: cell division; GTP binding
F;109-115/Region: tubulin/FtsZ GTP/GDP-binding (G-G-G-T-G-[ST]-G) motif

  Query Match             57.6%;  Score 510;  DB 2;  Length 398;
  Best Local Similarity   56.1%;  Pred. No. 3.1e-37;
  Matches  106;  Conservative   29;  Mismatches   42;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:|||| ||| :||   || || ::|||||: |:::| :||:   :|
Db         36 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDVGKGAAEESIDEIMEH 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL----------GILTVGVVSTPFRSEG 109
              ::||:|||| |||||| ||||||:|  |||                |||||||: ||  ||
Db         96 IKDSHMLFITAGMGGGTGTGAAPVIAKAAREARAAVKDRAPKEKKILTVGVVTKPFGFEG 155

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                |  :|  |::|| |||||||:|||||  :|:: ||  :||: ||:||  |::|||||
Db        156 VRRMPIAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFSDAFKLADNVLHIGIRGVTDL 215

Qy        170 IVRPGLINL 178
              :| ||||||
Db        216 MVMPGLINL 224


RESULT 7
H97108
cell division GTPase FtsZ [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: H97108
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: H97108
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK79659.1; PID:g15024657; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1693
C;Superfamily: cell division protein ftsZ

  Query Match             57.3%;  Score 507;  DB 2;  Length 373;
  Best Local Similarity   56.0%;  Pred. No. 5.4e-37;
  Matches   98;  Conservative   35;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |: ||||  ||| |||  | |  || :|  :|||||||: ||:|:::||: | .:| : ::
Db         35 LKNVEFIAINTDKQALALSQASQKIQIGDKLTKGLGAGANPEIGQKAAEESKDEISQAIK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:||| |||||| |||||||| :|: :|||||||: ||  ||  |   | :|:| |
Db         95 GADMVFITAGMGGGTGTGAAPVVAEIAKSMGILTVGVVTKPFPFEGRKRMLHAESGIKTL 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               : ||||: :||: |||: || ||::|||: ||||| :|||:|::|||  |||:||
Db        155 KERVDTLVTIPNERLLAIVDKKTTLVEAFKSADDVLRQGVQGISDLITIPGLVNL 209


RESULT 8
C70137
cell division protein ftsZ - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 26-Aug-1999
C;Accession: C70137; I40082; S21557
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: C70137
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-404 <KLE>
A;Cross-references: GB:AE001137; GB:AE000783; NID:g2688160; PIDN:AAC66649.1; PID:g268
A;Experimental source: strain B31
R;Old, I.G.; MacDougall, J.H.; Saint Girons, I.; Davidson, B.E.
FEMS Microbiol. Lett. 99, 245-250, 1992
A;Title: Mapping of genes on the linear chromosome of the bacterium Borrelia burgdorf
A;Reference number: I40082
A;Accession: I40082
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 34-119,'A' <RES>
A;Cross-references: EMBL:Z12164; NID:g49293; PID:g49294
A;Experimental source: strain 212
C;Comment: This cytoplasmic protein associates with the cell membrane just prior to c
C;Genetics:
A;Gene: ftsZ
C;Superfamily: cell division protein ftsZ
C;Keywords: cell division; GTP binding
F;120-126/Region: tubulin/FtsZ GTP/GDP-binding (G-G-G-T-G-[ST]-G) motif
F;221-228/Region: GTP-hydrolizing region

  Query Match             57.2%;  Score 506;  DB 2;  Length 404;
  Best Local Similarity   58.7%;  Pred. No. 7.2e-37;
  Matches  101;  Conservative   27;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          7 VEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDSN 66
              ||||||||| |||  |:|| || ||  :| ||||| |||:|: :||:    |:  |  ::
Db         52 VEFIVANTDLQALQTSIAPIKIALGAKVTAGLGAGGKPEIGQAAAEEDIDVIRNHLSGAD 111

Qy         67 MLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAKY 126
              |:||| |||||| ||||||:| ||:||||||||||: ||: ||| : |||  |:  | |
Db        112 MVFITAGMGGGTGTGAAPVIAQVAKELGILTVGVVTKPFKFEGPKKLRLAEQGINNLRKS 171

Qy        127 VDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              |||||::||| || : || ||: :||: |||||  ||:|:  ||:  | :|:
Db        172 VDTLIIIPNQKLLTVVDKRTTIKDAFKRADDVLRMGVQGIAGLIIEHGEVNI 223


RESULT 9
C87564
cell division protein FtsZ [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 17-May-2002
C;Accession: C87564
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87564
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-508 <STO>
A;Cross-references: GB:AE005673; NID:g13424103; PIDN:AAK24511.1; GSPDB:GN00148
```

```
C;Genetics:
A;Gene: CC2540
C;Superfamily: cell division protein ftsZ1

  Query Match             57.1%;  Score 505;  DB 2;  Length 508;
  Best Local Similarity   58.3%;  Pred. No. 1.2e-36;
  Matches  102;  Conservative   28;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ||||||:|||||  | |  :    :| ||  ||:|||||: ||:|  :||:   :|   |
Db         38 LEGVEFVVANTDAQQLQFAKTDRRIQLGVQITQGLGAGAHPEVGMSAAEESFPEIGEHLD 97

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|||  ||||||  |||||::|  ||| ||||||||: ||  || :| |||::|::||
Db         98 GAHMVFITAGMGGGTGTGAAPIIAKCARERGILTVGVVTKPFHFEGRHRMRLADSGIQEL 157

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :|||||||:|||||  :|:: ||  |||  || ||  ||: :|||:| ||||||
Db        158 QRYVDTLIVIPNQNLFRVANERTTFAEAFGMADQVLHSGVRSITDLMVLPGLINL 212


RESULT 10
I39848
cell division initiation protein (septum formation) FtsZ - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 19-Jul-1996 #sequence_revision 04-Oct-1996 #text_change 16-Jun-2000
C;Accession: I39848; A69628
R;Beall, B.; Lowe, M.; Lutkenhaus, J.
J. Bacteriol. 170, 4855-4864, 1988
A;Title: Cloning and characterization of Bacillus subtilis homologs of Escherichia coli
A;Reference number: I39846; MUID:89008108; PMID:3139638
A;Accession: I39848
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-382 <RES>
A;Cross-references: GB:M22630; NID:g142938; PIDN:AAA22457.1; PID:g142941
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69628
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-382 <KUN>
A;Cross-references: GB:Z99111; GB:AL009126; NID:g2633699; PIDN:CAB13402.1; PID:g2633900
A;Experimental source: strain 168
C;Comment: This cytoplasmic protein associates with the cell membrane just prior to cell
th vegetative and sporulating cell divisions.
C;Genetics:
A;Gene: ftsZ
C;Superfamily: cell division protein ftsZ
C;Keywords: cell division; GTP binding
F;106-112/Region: tubulin/FtsZ GTP/GDP-binding (G-G-G-T-G-[ST]-G) motif

  Query Match             56.3%;  Score 498;  DB 1;  Length 382;
  Best Local Similarity   53.7%;  Pred. No. 3.4e-36;
  Matches   95;  Conservative   37;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              ::::|||:|  ||| |||  | |  |: :|  :|:|||||: ||:|||::||: |  |:
Db         33 NEVQGVEYIAVNTDAQALNLSKAEVKMQIGAKLTRGLGAGANPEVGKKAAEESKEQIEEA 92

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
                |: ::|:|:|:| |||||| ||||||:| :|::|| ||||||: ||  ||  |   |  |:
Db         93 LKGADMVFVTAGMGGGTGTGAAPVIAQIAKDLGALTVGVVTRPFTFEGRKRQLQAAGGIS 152

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : : ||||||:||  :| : ||:| |||||| ||:|| :|||:|::||| ||||||
Db        153 AMKEAVDTLIVIPNDRILEIVDKNTPMLEAFREADNVLRQGVQGISDLIATPGLINL 209


RESULT 11
E86858
cell division protein FtsZ [imported] - Lactococcus lactis subsp. lactis (strain IL14
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: E86858
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: E86858
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-417 <STO>
A;Cross-references: GB:AE005176; PID:g12724901; PIDN:AAK05967.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ftsZ
C;Superfamily: cell division protein ftsZ

  Query Match             55.7%;  Score 493;  DB 2;  Length 417;
  Best Local Similarity   56.0%;  Pred. No. 1e-35;
  Matches   98;  Conservative   29;  Mismatches   48;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              : ||||| |||| |||  | |   | ||  :|:|||||::||:|||:||:    : : |:
Db         36 VSGVEFIAANTDVQALRSSKADTVIQLGPKLTRGLGAGAQPEVGKRAAEESAETVSQALE 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                |:|:||| |||||| |||||||:| :|:||| ||||||: ||  ||  |:  |  |:: |
Db         96 GSDMIFITAGMGGGTGTGAAPVIAQIAKELGALTVGVVTRPFGFEGSKRSYFATEGIEAL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                  |||::: | ||| : || | : || | ||:|| :|||:||||||  ||:|||
Db        156 RANVDTLLIISNNNLLEIVDKKTPLTEALREADNVLRQGVQGVTDLITNPGMINL 210


RESULT 12
S60765
cell division protein ftsZ - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 27-Apr-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C;Accession: S60765; T34952
R;McCormick, J.R.; Su, E.P.; Driks, A.; Losick, R.
Mol. Microbiol. 14, 243-254, 1994
A;Title: Growth and viability of Streptomyces coelicolor mutant for the cell division
A;Reference number: S60763; MUID:95131746; PMID:7830569
A;Accession: S60765
A;Molecule type: DNA
A;Residues: 1-399 <MCC>
A;Cross-references: EMBL:U10879; NID:g4204101; PIDN:AAD10533.1; PID:g527649
A;Experimental source: strain A3(2)
R;Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21563
A;Accession: T34952
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-399 <SAU>
A;Cross-references: EMBL:AL109663; PIDN:CAB51991.1; GSPDB:GN00070; SCOEDB:ftsZ
A;Experimental source: strain A3(2)
C;Comment: This cytoplasmic protein associates with the cell membrane just prior to c
C;Genetics:
```

```
A;Gene: ftsZ
A;Start codon: GTG
C;Superfamily: cell division protein ftsZ
C;Keywords: cell division; GTP binding
F;103-109/Region: tubulin/FtsZ GTP/GDP-binding (G-G-G-T-G-[ST]-G) motif
F;204-211/Region: GTP hydrolizing
F;383-391/Region: hydrophobic

  Query Match             54.9%;  Score 486;  DB 1;  Length 399;
  Best Local Similarity   53.1%;  Pred. No. 4.1e-35;
  Matches   93;  Conservative   38;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||||  ||| |||  | |  |: :|:::|:|||||: | :|:::||  : :|: :|:
Db         32 LKGVEFIAINTDAQALLMSDADVKLDVGRELTRGLGAGANPAVGRKAAEDHREEIEEVLK 91

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| | |||| || |||||::|| || ||:|||: ||  ||  |   |  |: ||
Db         92 GADMVFVTAGEGGGTGTGGAPVVANIARSLGALTIGVVTRPFTFEGRRRANQAEDGIAEL 151

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  ||:::|:  ::|:||: || ||| ||:|:|||| ||||||
Db        152 REEVDTLIVIPNDRLLSISDRQVSVLDAFKSADQVLLSGVQGITDLITTPGLINL 206


RESULT 13
JE0282
cell division protein ftsZ - Streptomyces collinus
C;Species: Streptomyces collinus
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 18-Aug-2000
C;Accession: JE0282
R;Zhulanova, E.; Mikulik, K.
Biochem. Biophys. Res. Commun. 249, 556-561, 1998
A;Title: Characterization of ftsZ gene and its protein product from Streptomyces collinu
A;Reference number: JE0282; MUID:98380301; PMID:9712736
A;Accession: JE0282
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-402 <ZHU>
A;Cross-references: GB:AF073487; NID:g3328123; PIDN:AAC33005.1; PID:g3328124
C;Genetics:
A;Gene: ftsZ
C;Function:
A;Description: required for septation and conversion of aerial mycelium into chains of s
C;Superfamily: cell division protein ftsZ
C;Keywords: cell division; GTP binding

  Query Match             54.9%;  Score 486;  DB 2;  Length 402;
  Best Local Similarity   53.1%;  Pred. No. 4.1e-35;
  Matches   93;  Conservative   38;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||||  ||| |||  | |  |: :|:::|:|||||: | :|:::||  : :|: :|:
Db         32 LKGVEFIAINTDAQALLMSDADVKLDVGRELTRGLGAGANPAVGRKAAEDHREEIEEVLK 91

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| | |||| || |||||::|| || ||:|||: ||  ||  |   |  |: ||
Db         92 GADMVFVTAGEGGGTGTGGAPVVANIARSLGALTIGVVTRPFTFEGRRRANQAEDGIAEL 151

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  ||:::|:  ::|:||: || ||| ||:|:|||| ||||||
Db        152 REEVDTLIVIPNDRLLSISDRQVSVLDAFKSADQVLLSGVQGITDLITTPGLINL 206


RESULT 14
B70579
probable cell division protein FtsZ - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: B70579
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: B70579
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <COL>
A;Cross-references: GB:Z95388; GB:AL123456; NID:g3261759; PIDN:CAB08643.1; PID:g21043
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: ftsZ
C;Superfamily: cell division protein ftsZ

  Query Match             54.8%;  Score 485;  DB 2;  Length 379;
  Best Local Similarity   54.3%;  Pred. No. 4.7e-35;
  Matches   95;  Conservative   35;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||||  ||| |||  | |  |: :|:| |:|||||: ||:|:::||  | :|: :|:
Db         32 LKGVEFIAINTDAQALLMSDADVKLDVGRDSTRGLGAGADPEVGRKAAEDAKDEIEELLR 91

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| | |||| || ||||||:||:|| ||||||: ||  ||  |:  |  |:  |
Db         92 GADMVFVTAGEGGGTGTGGAPVVASIARKLGALTVGVVTRPFSFEGKRRSNQAENGIAAL 151

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              :  |||||:||  || : | : ::::||| ||:||| ||:|:|||| |||||:
Db        152 RESCDTLIVIPNDRLLQMGDAAVSLMDAFRSADEVLLNGVQGITDLITTPGLINV 206


RESULT 15
C97610
cell division protein ftsz [imported] - Agrobacterium tumefaciens (strain C58, Cereon
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C;Accession: C97610
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A;Reference number: A97359; PMID:11743194
A;Accession: C97610
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-583 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK87836.1; PID:g15157218; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_3784
A;Map position: circular chromosome
C;Superfamily: cell division protein ftsZ1

  Query Match             54.5%;  Score 482;  DB 2;  Length 583;
  Best Local Similarity   57.1%;  Pred. No. 1.4e-34;
  Matches  100;  Conservative   26;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:||:|:||||| |||  : |   | || ::|:||||||:||:|: :||:   :|   |
Db         38 LQGVDFVVANTDAQALTMTKADRVIQLGVNVTEGLGAGSQPEVGRAAAEECIDEIIDHLN 97

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::| |:| |||||| |||||||  ||  |||||||: ||  ||  | |||  |::||
Db         98 GTHMCFVTAGMGGGTGTGAAPVVAQAARNKGILTVGVVTKPFHFEGGRRMRLAEQGIEEL 157

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              | ||||||:|||||  :|:  ||  :||  || ||  ||  :|||:|: |||||
Db        158 QKSVDTLIVIPNQNLFRIANDKTTFADAFAMADQVLYSGVACITDLMVKEGLINL 212


Search completed: June  2, 2003, 08:42:15
Job time : 49 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2003, 06:02:36 ; Search time 22 Seconds
(without alignments)
335.581 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	524	59.2	452	1	FTSZ_RICCN	Q92gv7 rickettsia
2	518	58.5	452	1	FTSZ_RICPR	Q9zcq3 rickettsia
3	511	57.7	382	1	FTSZ_BACHD	Q9k9t7 bacillus ha
4	510	57.6	398	1	FTSZ_WOLSP	P45485 wolbachia s
5	506	57.2	399	1	FTSZ_BORBU	P45483 borrelia bu
6	505	57.1	508	1	FTSZ_CAUCR	P52976 caulobacter
7	498	56.3	382	1	FTSZ_BACSU	P17865 bacillus su
8	494	55.8	592	1	FTSZ_BARBA	O31314 bartonella
9	492	55.6	413	1	FTSZ_ENTHR	O08458 enterococcu
10	489	55.3	412	1	FTSZ_ENTFA	O08439 enterococcu
11	486	54.9	386	1	FTSZ_MYCKA	Q9kh25 mycobacteri
12	486	54.9	399	1	FTSZ_STRCO	P45500 streptomyce
13	485	54.8	379	1	FTSZ_MYCTU	O08378 mycobacteri
14	485	54.8	407	1	FTSZ_STRGR	P45501 streptomyce
15	482	54.5	583	1	FTSZ_AGRTU	O30992 agrobacteri
16	481	54.4	379	1	FTSZ_MYCLE	Q9cce4 mycobacteri
17	480	54.2	390	1	FTSZ_STAAM	P45498 staphylococ
18	477	53.9	381	1	FTSZ_FRATU	Q9zaw3 francisella
19	475	53.7	590	1	FTZ1_RHIME	P30327 rhizobium m
20	471	53.2	400	1	FTSZ_PSEPU	Q59692 pseudomonas
21	470	53.1	442	1	FTSZ_CORGL	P94337 corynebacte
22	468	52.9	428	1	FTSZ_ANASP	P45482 anabaena sp
23	467	52.8	430	1	FTSZ_SYNY3	P73456 synechocyst
24	465	52.5	384	1	FTSZ_WIGGL	Q9ala3 wiggleswort
25	465	52.5	418	1	FTSZ_TREPA	O83405 treponema p
26	464	52.4	394	1	FTSZ_PSEAE	P47204 pseudomonas
27	461	52.1	384	1	FTSZ_BUCAI	P57308 buchnera ap
28	459	51.9	346	1	FTZ2_RHIME	P45484 rhizobium m
29	458	51.8	170	1	FTSZ_MICLU	P45499 micrococcus
30	453	51.2	384	1	FTSZ_BUCAP	O51929 buchnera ap
31	453	51.2	394	1	FTSZ_AZOVI	P77817 azotobacter
32	452	51.1	383	1	FTSZ_ECOLI	P06138 escherichia
33	450	50.8	386	1	FTSZ_SODGL	Q9ala4 sodalis glo
34	447	50.5	421	1	FTSZ_HAEIN	P45069 haemophilus
35	446	50.4	351	1	FTSZ_THEMA	O08398 thermotoga
36	445	50.3	315	1	FTSZ_WOLPI	P50907 wolbachia p
37	440	49.7	392	1	FTSZ_NEIMA	Q51130 neisseria m
38	435	49.2	433	1	FTSZ_ARATH	Q42545 arabidopsis
39	434	49.0	392	1	FTSZ_NEIGO	P72079 neisseria g
40	431.5	48.8	366	1	FTZ1_PYRWO	Q52630 pyrococcus
41	431.5	48.8	372	1	FTZ1_PYRHO	O57776 pyrococcus
42	424.5	48.0	372	1	FTZ1_PYRAB	Q9v2s0 pyrococcus
43	411	46.4	390	1	FTSZ_MYCPU	Q50318 mycoplasma
44	407.5	46.0	368	1	FTZ1_ARCFU	O29715 archaeoglob
45	407	46.0	367	1	FTSZ_AQUAE	O66809 aquifex aeo

ALIGNMENTS

RESULT 1
FTSZ_RICCN
ID FTSZ_RICCN STANDARD; PRT; 452 AA.
AC Q92GV7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division protein ftsZ.
GN FTSZ OR RC1015.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
CC -!- FUNCTION: This protein is essential to the cell-division process.
CC It seems to assemble into a dynamic ring on the inner surface of
CC the cytoplasmic membrane at the place where division will occur,
CC and the formation of the ring is the signal for septation to
CC begin. Binds to and hydrolyzes GTP (By similarity).
CC -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC of the cytoplasmic membrane (By similarity).
CC -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC ----------

CC ----------
DR EMBL; AE008653; AAL03553.1; -.
DR InterPro; IPR000158; FtsZ.
DR InterPro; IPR003008; Tubulin_FtsZ.
DR Pfam; PF00091; tubulin; 1.
DR TIGRFAMs; TIGR00065; ftsZ; 1.
DR PROSITE; PS01134; FTSZ_1; 1.
DR PROSITE; PS01135; FTSZ_2; 1.
KW Cell division; Septation; GTP-binding; Complete proteome.
FT NP_BIND 107 115 GTP (POTENTIAL).
SQ SEQUENCE 452 AA; 48400 MW; B326101A609A43EA CRC64;

Query Match 59.2%; Score 524; DB 1; Length 452;
Best Local Similarity 56.2%; Pred. No. 1.8e-37;
Matches 100; Conservative 38; Mismatches 40; Indels 0; Gaps 0;

Qy 1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60

```
              :: |:|  |:|||||| |:|  ||  :|| ||   |:||||||: ||:|  :|:: : :|:
Db         35 SANLQGANFVVANTDAQSLEHSLCTNKIQLGVSTTRGLGAGASPEVGALAAQESESEIRN 94

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
              |::|||:||| |||||| ||:|||:| :|:|||||||||: ||  || :| : |: |:
Db         95 YLENSNMVFITAGMGGGTGTGSAPVIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGL 154

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || ::||||||:|||||  :|:: ||  :||: |||||  ||:||||||:: ||||||
Db        155 IELQQFVDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINL 212


RESULT 2
FTSZ_RICPR
ID   FTSZ_RICPR     STANDARD;      PRT;   452 AA.
AC   Q9ZCQ3;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ OR RP666.
OS   Rickettsia prowazekii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=782;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Madrid E;
RX   MEDLINE=99039499; PubMed=9823893;
RA   Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA   Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA   Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT   "The genome sequence of Rickettsia prowazekii and the origin of
RT   mitochondria.";
RL   Nature 396:133-140(1998).
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ235272; CAA15104.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding; Complete proteome.
FT   NP_BIND     107    115       GTP (POTENTIAL).
SQ   SEQUENCE   452 AA;  48806 MW;  D6AC69F28B3343AE CRC64;

  Query Match             58.5%;  Score 518;  DB 1;  Length 452;
  Best Local Similarity   55.9%;  Pred. No. 5.8e-37;
  Matches   99;  Conservative   38;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              : |:|  |:|||||| |:|  ||  :|| ||   |:||||||: ||:|  :|:: : :|:
Db         36 ANLQGANFVVANTDAQSLEHSLCINKIQLGVSTTRGLGAGASPEVGALAAQESENEIRSS 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
              |::|||:||| |||||| ||:|||:| :|:|||||||||: ||  || :| : |: |:
Db         96 LENSNMVFITAGMGGGTGTGSAPIIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGLI 155

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || ::||||||:|||||  :|:: ||  :||: |||||  ||:||||||:: ||||||
Db        156 ELQQFVDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINL 212


RESULT 3
FTSZ_BACHD
ID   FTSZ_BACHD     STANDARD;      PRT;   382 AA.
AC   Q9K9T7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ OR BH2558.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP001515; BAB06277.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding; Complete proteome.
FT   NP_BIND     104    112       GTP (POTENTIAL).
SQ   SEQUENCE   382 AA;  40583 MW;  9F309B5AE4E2AC1E CRC64;

  Query Match             57.7%;  Score 511;  DB 1;  Length 382;
  Best Local Similarity   58.3%;  Pred. No. 1.9e-36;
  Matches  102;  Conservative   30;  Mismatches   43;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:||:||  ||| |||  | |  |: ||  :|:|||||: ||:||::||: :  |:  ||
Db         35 LQGVDFISVNTDAQALHLSKAEVKLQLGGKLTRGLGAGANPEIGKKAAEESREQIEEALQ 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ::|:||| |||||| ||||||:| ||:|:| ||||||: ||  ||  |:  | ||:  |
```

```
Db         95 GADMVFITAGMGGGTGTGAAPVIAEVAKEIGALTVGVVTRPFTFEGRKRSTQAAAGIAAL 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  || : ||:| |||||| ||:|| :||:|::|||  ||||||
Db        155 KEKVDTLIVIPNDRLLEIVDKNTPMLEAFREADNVLRQGVQGISDLIATPGLINL 209


RESULT 4
FTSZ_WOLSP
ID   FTSZ_WOLSP     STANDARD;      PRT;   398 AA.
AC   P45485;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Wolbachia sp.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Wolbachieae; Wolbachia.
OX   NCBI_TaxID=956;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93360900; PubMed=7689140;
RA   Holden P.R., Brookfield J.F.Y., Jones P.;
RT   "Cloning and characterization of an ftsZ homologue from a bacterial
RT   symbiont of Drosophila melanogaster.";
RL   Mol. Gen. Genet. 240:213-220(1993).
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X71906; CAA50724.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding.
FT   NP_BIND     107    115       GTP (POTENTIAL).
SQ   SEQUENCE   398 AA;  42497 MW;  4D4747719C67E84D CRC64;

  Query Match             57.6%;  Score 510;  DB 1;  Length 398;
  Best Local Similarity   56.1%;  Pred. No. 2.4e-36;
  Matches  106;  Conservative   29;  Mismatches   42;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:||||| ||| :||   || || ::|||||||: |::|| :||:   :|
Db         36 SNLQGVNFVVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDVGKGAAEESIDEIMEH 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
              ::||:||||| |||||| ||||||:|  |||           |||||||: ||  ||
Db         96 IKDSHMLFITAGMGGGTGTGAAPVIAKAAREARAAVKDRAPKEKKILTVGVVTKPFGFEG 155

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                |  :|  |::|| |||||||:|||||  :|:: ||  :||: ||:||  |::|||||
Db        156 VRRMPIAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFSDAFKLADNVLHIGIRGVTDL 215

Qy        170 IVRPGLINL 178
                 :| ||||||
Db        216 MVMPGLINL 224


RESULT 5
FTSZ_BORBU
ID   FTSZ_BORBU     STANDARD;      PRT;   399 AA.
AC   P45483; Q59183;
DT   01-NOV-1995 (Rel. 32, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ OR BB0299.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RA   Dunn J.J., Butler-Loffredo L., Kieleczawa J., Medalle J., Luft B.J.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=212;
RA   Ge Y., Old I.G., Saint-Girons I., Charon N.W.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 29-115 FROM N.A.
RC   STRAIN=212;
RX   MEDLINE=93146383; PubMed=1490605;
RA   Old I.G., Macdougall J.H., Saint-Girons I., Davidson B.E.;
RT   "Mapping of genes on the linear chromosome of the bacterium Borrelia
RT   burgdorferi: possible locations for its origin of replication.";
RL   FEMS Microbiol. Lett. 78:245-250(1992).
RN   [5]
RP   SEQUENCE OF 1-115 FROM N.A.
RC   STRAIN=212;
RA   Old I.G.;
RL   Submitted (JUN-1992) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; U43739; AAA85622.1; ALT_INIT.
DR   EMBL; AE001137; AAC66649.1; ALT_INIT.
DR   EMBL; X96685; CAA65464.1; ALT_INIT.
DR   EMBL; L76303; AAB51402.1; -.
DR   EMBL; Z12164; CAA78156.1; ALT_INIT.
DR   HSSP; Q57816; 1FSZ.
DR   TIGR; BB0299; -.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding; Complete proteome.
FT   NP_BIND     113    121       GTP (POTENTIAL).
FT   CONFLICT    115    115       G -> A (IN REF. 4 AND 5).
FT   CONFLICT    249    250       AA -> RR (IN REF. 3).
FT   CONFLICT    336    336       T -> A (IN REF. 3).
SQ   SEQUENCE   399 AA;  42398 MW;  E808E336343EE583 CRC64;

  Query Match             57.2%;  Score 506;  DB 1;  Length 399;
  Best Local Similarity   58.7%;  Pred. No. 5.3e-36;
  Matches  101;  Conservative   27;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          7 VEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQDSN 66
              |||||||||| |||  |:|| || ||  :| ||||| |||:|: :||:    |:  |  ::
Db         47 VEFIVANTDLQALQTSIAPIKIALGAKVTAGLGAGGKPEIGQAAAEEDIDVIRNHLSGAD 106

Qy         67 MLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKELAKY 126
              |:||| |||||| ||||||:| ||:||||||||||: ||: ||| : |||  |:  | |
Db        107 MVFITAGMGGGTGTGAAPVIAQVAKELGILTVGVVTKPFKFEGPKKLRLAEQGINNLRKS 166

Qy        127 VDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              |||||::||| || : || ||: :||: |||||  ||:|:  ||:  | :|:
Db        167 VDTLIIIPNQKLLTVVDKRTTIKDAFKRADDVLRMGVQGIAGLIIEHGEVNI 218


RESULT 6
FTSZ_CAUCR
ID   FTSZ_CAUCR     STANDARD;      PRT;   508 AA.
AC   P52976;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ OR CC2540.
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC   Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CB15N / NA1000;
RX   MEDLINE=96270535; PubMed=8692812;
RA   Quardokus E., Din N., Brun Y.V.;
RT   "Cell cycle regulation and cell type-specific localization of the
RT   FtsZ division initiation protein in Caulobacter.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:6314-6319(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA   Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA   Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U40273; AAC44223.1; -.
DR   EMBL; AE005922; AAK24511.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   TIGR; CC2540; -.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding; Complete proteome.
FT   NP_BIND     107    115       GTP (POTENTIAL).
FT   CONFLICT    236    236       G -> A (IN REF. 1).
SQ   SEQUENCE   508 AA;  54207 MW;  4E817C099F58C4CE CRC64;

  Query Match             57.1%;  Score 505;  DB 1;  Length 508;
  Best Local Similarity   58.3%;  Pred. No. 8.5e-36;
  Matches  102;  Conservative   28;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ||||||:||||| | |  :    :| ||  ||:|||||: ||:|  :||:   :|   |
Db         38 LEGVEFVVANTDAQQLQFAKTDRRIQLGVQITQGLGAGAHPEVGMSAAEESFPEIGEHLD 97

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:||| |||||| |||||::|  ||| ||||||||: ||  || :| |||::|::||
Db         98 GAHMVFITAGMGGGTGTGAAPIIAKCARERGILTVGVVTKPFHFEGRHRMRLADSGIQEL 157

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               :|||||||:||||| :|:: ||  |||  || ||  ||: :|||:| ||||||
Db        158 QRYVDTLIVIPNQNLFRVANERTTFAEAFGMADQVLHSGVRSITDLMVLPGLINL 212


RESULT 7
FTSZ_BACSU
ID   FTSZ_BACSU     STANDARD;      PRT;   382 AA.
AC   P17865;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89008108; PubMed=3139638;
RA   Beall B., Lowe M., Lutkenhaus J.;
RT   "Cloning and characterization of Bacillus subtilis homologs of
RT   Escherichia coli cell division genes ftsZ and ftsA.";
RL   J. Bacteriol. 170:4855-4864(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
```

```
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [3]
RP   SEQUENCE OF 371-382 FROM N.A.
RX   MEDLINE=90216713; PubMed=2108961;
RA   Wu X.-C., Nathoo S., Pang A.S.-H., Carne T., Wang S.-L.;
RT   "Cloning, genetic organization, and characterization of a structural
RT   gene encoding bacillopeptidase F from Bacillus subtilis.";
RL   J. Biol. Chem. 265:6845-6850(1990).
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22630; AAA22457.1; -.
DR   EMBL; Z99111; CAB13402.1; -.
DR   EMBL; J05400; AAA83361.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   SubtiList; BG10232; ftsZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding; Complete proteome.
FT   NP_BIND     104    112       GTP (POTENTIAL).
SQ   SEQUENCE   382 AA;  40355 MW;  D1E908DED2734CBE CRC64;
```

```
  Query Match             56.3%;  Score 498;  DB 1;  Length 382;
  Best Local Similarity   53.7%;  Pred. No. 2.5e-35;
  Matches   95;  Conservative   37;  Mismatches   45;  Indels    0;  Gaps    0;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              ::::|||:|  ||| |||  | |  |: :|  :|:|||||: ||:||::||: |  |:
Db         33 NEVQGVEYIAVNTDAQALNLSKAEVKMQIGAKLTRGLGAGANPEVGKKAAEESKEQIEEA 92

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
              |: ::|:|:| |||||| ||||||:| :|::|| ||||||: ||  ||  |   |  |:
Db         93 LKGADMVFVTAGMGGGTGTGAAPVIAQIAKDLGALTVGVVTRPFTFEGRKRQLQAAGGIS 152

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
                : : ||||||:||  :| : ||:| |||||| ||:|| :||:|::|||  ||||||
Db        153 AMKEAVDTLIVIPNDRILEIVDKNTPMLEAFREADNVLRQGVQGISDLIATPGLINL 209


RESULT 8
FTSZ_BARBA
ID   FTSZ_BARBA     STANDARD;      PRT;   592 AA.
AC   O31314;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ (75 kDa antigen).
GN   FTSZ.
OS   Bartonella bacilliformis.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Bartonellaceae; Bartonella.
OX   NCBI_TaxID=774;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KC584;
RX   MEDLINE=97369823; PubMed=9226264;
RA   Padmalayam I., Anderson B., Kron M., Kelly T., Baumstark B.;
RT   "The 75-kilodalton antigen of Bartonella bacilliformis is a
RT   structural homolog of the cell division protein ftsZ.";
RL   J. Bacteriol. 179:4545-4552(1997).
RN   [2]
RP   REVISIONS.
RA   Kelly T.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF007266; AAC15082.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding.
FT   NP_BIND     107    115       GTP (POTENTIAL).
SQ   SEQUENCE   592 AA;  63503 MW;  C32007DADCD2D75B CRC64;
```

```
  Query Match             55.8%;  Score 494;  DB 1;  Length 592;
  Best Local Similarity   57.7%;  Pred. No. 8.7e-35;
  Matches  101;  Conservative   28;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:||:|:|||||  |||  | |   | ||  :|:||||||: ||:|: :|::   :|   |
Db         38 LQGVDFVVANTDAQALAMSKAERVIQLGAAVTEGLGAGALPEVGQAAADECIDEIIDHLA 97

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ||:|:||| ||||||  ||||||||  ||| ||||||||: ||: ||  | : | ||::||
Db         98 DSHMVFITAGMGGGTGTGAAPVVARAAREKGILTVGVVTKPFQFEGARRMKTAEAGIEEL 157

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              | ||||||:|||||  :|:: ||  :||  || ||  ||  :|||::: |||||
Db        158 QKSVDTLIVIPNQNLFRIANEKTTFADAFAMADQVLYSGVASITDLMIKEGLINL 212


RESULT 9
FTSZ_ENTHR
ID   FTSZ_ENTHR     STANDARD;      PRT;   413 AA.
AC   O08458;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Enterococcus hirae.
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S185;
RX   MEDLINE=99449055; PubMed=10520745;
RA   Duez C., Thamm I., Sapunaric F., Coyette J., Ghuysen J.-M.;
RT   "The division and cell wall gene cluster of Enterococcus hirae S185.";
RL   DNA Seq. 9:149-161(1998).
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y13922; CAA74240.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding.
FT   NP_BIND     105    113       GTP (POTENTIAL).
SQ   SEQUENCE   413 AA;  44268 MW;  581AE9F85702A097 CRC64;

  Query Match             55.6%;  Score 492;  DB 1;  Length 413;
  Best Local Similarity   55.4%;  Pred. No. 8.7e-35;
  Matches   97;  Conservative   31;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  |||| |||  | |   | ||   |:||||||:||:|:::||: :  ::  |
Db         36 VKGVEFITANTDVQALKNSKAETVIQLGPKYTRGLGAGSQPEVGQKAAEESEQALREALD 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ::|:|||  ||||||  |||||:|| :|:||| ||||||: ||  ||| | | |  |:  |
Db         96 GADMIFITAGMGGGTGTGAAPIVAGIAKELGALTVGVVTRPFTFEGPKRGRFAAEGIARL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        156 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 210


RESULT 10
FTSZ_ENTFA
ID   FTSZ_ENTFA     STANDARD;      PRT;   412 AA.
AC   O08439;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A24836;
RA   Pucci M.J., Thanassi J.A., Discotto L.F., Kessler R.E.,
RA   Dougherty T.J.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U94707; AAC45639.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; FALSE_NEG.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding.
FT   NP_BIND     105    113       GTP (POTENTIAL).
SQ   SEQUENCE   412 AA;  44355 MW;  42EA85BAA70EF51F CRC64;

  Query Match             55.3%;  Score 489;  DB 1;  Length 412;
  Best Local Similarity   56.6%;  Pred. No. 1.6e-34;
  Matches   99;  Conservative   28;  Mismatches   48;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  |||| |||  | |   | ||   |: |||||:||:|:::||: :  |  ||
Db         36 VKGVEFITANTDVQALKHSKAETVIQLGPKYTRNLGAGSQPEVGQKAAEESEQVISESLQ 95

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
              ::|:|||  ||||||  |||||||| :|:||| ||||||: ||  ||| | | |  |:  |
Db         96 GADMIFITAGMGGGTGTGAAPVVAKIAKELGALTVGVVTRPFSFEGPKRGRFAAEGIALL 155

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
```

```
              : ||||::: |  || : || | |||||| ||:|| :||:|::|||  || :||
Db        156 KENVDTLLIISNNRLLEVVDKKTPMLEAFREADNVLRQGVQGISDLITAPGYVNL 210


RESULT 11
FTSZ_MYCKA
ID   FTSZ_MYCKA     STANDARD;      PRT;   386 AA.
AC   Q9KH25;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Mycobacterium kansasii.
OC   Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1768;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ramanujam S., Ajitkumar P.;
RT   "ftsZ gene of M. kansasii.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF273451; AAF78784.2; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding.
FT   NP_BIND     101    109       GTP (POTENTIAL).
SQ   SEQUENCE   386 AA;  39051 MW;  215DE0B814EED593 CRC64;

  Query Match           54.9%;  Score 486;  DB 1;  Length 386;
  Best Local Similarity   54.9%;  Pred. No. 2.6e-34;
  Matches   96;  Conservative   33;  Mismatches   46;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||||  ||| |||  | |  |: :|:| |:|||||: ||:|: :||  | ||: :|:
Db         32 LKGVEFIAINTDAQALLMSDADVKLDVGRDSTRGLGAGADPEVGRXAAEDAKDDIEELLR 91

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:|:| | |||| || ||||||:||:|| |||||||: ||  ||  |:  |  |:: |
Db         92 GADMVFVTAGEGGGTGTGGAPVVASIARKLGALTVGVVTRPFSFEGKRRSNQAENGIQAL 151

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               :  |||||:||  || : | : ::::||| ||:||| || |:|||| |||||:
Db        152 RESCDTLIVIPNDRLLQMGDAAVSLMDAFRSADEVLLNGVXGITDLITTPGLINV 206


RESULT 12
FTSZ_STRCO
ID   FTSZ_STRCO     STANDARD;      PRT;   399 AA.
AC   P45500;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ OR SCO2082 OR SC4A10.15c.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=95131746; PubMed=7830569;
RA   McCormick J.R., Su E.P., Driks A., Losick R.;
RT   "Growth and viability of Streptomyces coelicolor mutant for the cell
RT   division gene ftsZ.";
RL   Mol. Microbiol. 14:243-254(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U10879; AAD10533.1; -.
DR   EMBL; AL109663; CAB51991.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding; Complete proteome.
FT   NP_BIND     101    109       GTP (POTENTIAL).
SQ   SEQUENCE   399 AA;  41095 MW;  EAD52B04CFBB4D39 CRC64;

  Query Match           54.9%;  Score 486;  DB 1;  Length 399;
  Best Local Similarity   53.1%;  Pred. No. 2.7e-34;
  Matches   93;  Conservative   38;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||||  ||| |||  | |  |: :|:::|:|||||: | :|:::||  : :|: :|:
Db         32 LKGVEFIAINTDAQALLMSDADVKLDVGRELTRGLGAGANPAVGRKAAEDHREEIEEVLK 91
```

```
Qy          64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:|:| | |||| || |||||::|| || ||:|||: ||  ||  |   |  |: ||
Db          92 GADMVFVTAGEGGGTGTGGAPVVANIARSLGALTIGVVTRPFTFEGRRRANQAEDGIAEL 151

Qy         124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               : ||||||:||  ||:::|:  ::|:||: || ||| ||:|:|||| ||||||
Db         152 REEVDTLIVIPNDRLLSISDRQVSVLDAFKSADQVLLSGVQGITDLITTPGLINL 206


RESULT 13
FTSZ_MYCTU
ID   FTSZ_MYCTU     STANDARD;      PRT;   379 AA.
AC   O08378;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ OR RV2150C OR MT2209 OR MTCY270.18.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White.O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z95388; CAB08643.1; -.
DR   EMBL; AE007068; AAK46493.1; ALT_INIT.
DR   HSSP; Q57816; 1FSZ.
DR   TIGR; MT2209; -.
DR   TubercuList; Rv2150c; -.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding; Complete proteome.
FT   NP_BIND     101    109       GTP (POTENTIAL).
SQ   SEQUENCE   379 AA;  38756 MW;  3F580353078788A9 CRC64;

  Query Match             54.8%;  Score 485;  DB 1;  Length 379;
  Best Local Similarity   54.3%;  Pred. No. 3.1e-34;
  Matches   95;  Conservative   35;  Mismatches   45;  Indels    0;  Gaps    0;

Qy           4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
               |:|||||  ||| |||  | |  |: :|:| |:|||||: ||:|:::||  | :|: :|:
Db          32 LKGVEFIAINTDAQALLMSDADVKLDVGRDSTRGLGAGADPEVGRKAAEDAKDEIEELLR 91

Qy          64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
               ::|:|:| | |||| || ||||||:||:|| ||||||: ||  ||  |:  |  |: |
Db          92 GADMVFVTAGEGGGTGTGGAPVVASIARKLGALTVGVVTRPFSFEGKRRSNQAENGIAAL 151

Qy         124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               :  |||||:||  || : | : ::::||| ||:||| ||:|:|||| |||||:
Db         152 RESCDTLIVIPNDRLLQMGDAAVSLMDAFRSADEVLLNGVQGITDLITTPGLINV 206


RESULT 14
FTSZ_STRGR
ID   FTSZ_STRGR     STANDARD;      PRT;   407 AA.
AC   P45501;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Streptomyces griseus.
OC   Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1911;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRRL B2682;
RX   MEDLINE=94374704; PubMed=8088545;
RA   Dharmatilake A.J., Kendrick K.E.;
RT   "Expression of the division-controlling gene ftsZ during growth and
RT   sporulation of the filamentous bacterium Streptomyces griseus.";
RL   Gene 147:21-28(1994).
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U07344; AAA56889.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 2, 2003, 06:27:41 ; Search time 72 Seconds
(without alignments)
509.394 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	547	61.8	253	2	Q8RMK5	Q8rmk5 azospirillu
2	527	59.5	336	2	Q9RNN2	Q9rnn2 zymomonas m
3	527	59.5	401	10	Q9M7M6	Q9m7m6 mallomonas
4	525	59.3	421	2	Q9AQ38	Q9aq38 ehrlichia c
5	524	59.2	452	2	Q9AQ36	Q9aq36 rickettsia
6	521	58.9	400	2	Q9AQ37	Q9aq37 anaplasma p
7	520	58.8	357	16	Q8R9H2	Q8r9h2 thermoanaer
8	520	58.8	370	2	O85474	O85474 clostridium
9	517	58.4	231	2	P77997	P77997 wolbachia s
10	517	58.4	289	2	O08390	O08390 wolbachia s
11	515	58.2	289	2	O08392	O08392 wolbachia s
12	514	58.1	289	2	O08471	O08471 wolbachia s
13	514	58.1	289	2	O08389	O08389 wolbachia s
14	514	58.1	289	2	O08391	O08391 wolbachia s
15	510	57.6	391	16	Q8Y5M5	Q8y5m5 listeria mo
16	510	57.6	392	16	Q929Y5	Q929y5 listeria in
17	507	57.3	373	16	Q97IE9	Q97ie9 clostridium
18	503	56.8	581	2	O69074	O69074 bartonella
19	499	56.4	590	2	O69075	O69075 bartonella
20	497	56.2	350	2	Q9F7C6	Q9f7c6 wolbachia e
21	496	56.0	581	2	Q9X6M9	Q9x6m9 bartonella
22	495	55.9	381	16	Q8XJJ1	Q8xjj1 clostridium
23	493	55.7	372	2	O85475	O85475 clostridium
24	493	55.7	417	16	Q9CEH2	Q9ceh2 lactococcus
25	491	55.5	351	2	Q9L7Q8	Q917q8 wolbachia s
26	491	55.5	351	2	Q9L6V2	Q916v2 wolbachia e
27	490	55.4	407	10	Q9SSV6	Q9ssv6 cyanidiosch
28	489	55.3	392	2	Q9F1M6	Q9f1m6 shewanella
29	488	55.1	419	2	Q9ZAJ1	Q9zaj1 lactococcus
30	487	55.0	559	16	Q98KB9	Q98kb9 rhizobium l
31	486.5	55.0	318	2	O86036	O86036 wolbachia e
32	486	54.9	290	2	O85473	O85473 epulopisciu
33	486	54.9	402	2	O85717	O85717 streptomyce
34	485	54.8	348	2	O86037	O86037 wolbachia e
35	482	54.5	583	16	Q8UDN5	Q8udn5 agrobacteri
36	481	54.4	373	2	Q9RME1	Q9rme1 acholeplasm
37	476	53.8	343	16	Q986Q2	Q986q2 rhizobium l
38	476	53.8	538	16	Q8YI60	Q8yi60 brucella me
39	475	53.7	566	2	Q93TA1	Q93ta1 brucella ab
40	472	53.3	419	2	Q9ZHB9	Q9zhb9 streptococc
41	472	53.3	419	16	Q97PF9	Q97pf9 streptococc
42	471	53.2	336	2	Q9Z3B8	Q9z3b8 wolbachia s
43	469	53.0	331	2	Q9EVS3	Q9evs3 wolbachia s
44	468	52.9	331	2	Q9EVS4	Q9evs4 wolbachia s
45	468	52.9	334	2	Q9EVS0	Q9evs0 wolbachia s

ALIGNMENTS

RESULT 1
Q8RMK5
ID Q8RMK5 PRELIMINARY; PRT; 253 AA.
AC Q8RMK5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cell division protein FtsZ (Fragment).
GN FTSZ.
OS Azospirillum brasilense.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC Azospirillum.
OX NCBI_TaxID=192;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD;
RA Jofre E.C., Lagares A., Mori G.B.;
RT "A ddlB (D-alanine-D-alanine ligase) Tn5 mutant of Azospirillum
RT brasilense is altered in exopolysaccharide production, salt tolerance,
RT and maize root colonization.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF492457; AAM11652.1; -.
FT NON_TER 253 253
SQ SEQUENCE 253 AA; 26638 MW; 13B8B38B0AF4702D CRC64;

Query Match 61.8%; Score 547; DB 2; Length 253;
Best Local Similarity 60.5%; Pred. No. 2.9e-41;
Matches 107; Conservative 30; Mismatches 40; Indels 0; Gaps 0;

```
Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | ||||:|:| ||| ||| ||   :: ||  :|:|||||||::|: |||:|  :|
Db         36 SNLEGVDFVVGNTDAQALKGSLCEKRVQLGTTMTRGLGAGSKPDVGRASAEEQLEEIIGH 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
              |: :||:||| |||||| ||||||:|  ||| |:||||||: ||  || :| ||| :|:
Db         96 LEGANMVFITAGMGGGTGTGAAPVIARAARERGLLTVGVVTKPFHFEGAHRMRLAESGIA 155

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || :|||||::|||||  :|:: ||  :||: |||||  ||:||||:| ||||||
```

```
Db        156 ELQQYVDTLIIIPNQNLFRIANEKTTFADAFKMADDVLHSGVRGVTDLMVMPGLINL 212


RESULT 2
Q9RNN2
ID   Q9RNN2       PRELIMINARY;      PRT;   336 AA.
AC   Q9RNN2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Zymomonas mobilis.
OC   Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC   Zymomonas.
OX   NCBI_TaxID=542;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ZM4;
RA   Um H.W., Kang H.S.;
RT   "Sequence analysis of 41E10 fosmid clone of Zymomonas mobilis.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; AF179611; AAD53930.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
SQ   SEQUENCE   336 AA;  35725 MW;  8CEFF1F7150436C7 CRC64;

  Query Match             59.5%;  Score 527;  DB 2;  Length 336;
  Best Local Similarity   59.6%;  Pred. No. 2.6e-39;
  Matches  106;  Conservative   29;  Mismatches   43;  Indels    0;  Gaps    0;

Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              || ::||:||||||| |||  | |  :| ||   |:||||||:||:|| :||:    ||
Db         38 ASGVQGVDFIVANTDAQALNISPAEQRIQLGPTTTQGLGAGSRPEVGKAAAEETIEQIQE 97

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
               |: : | ||  |||||| ||||||:| |||:.||||||||: ||  ||  | | | :|:
Db         98 ALEGARMCFIAAGMGGGTGTGAAPVIAKVARDRGILTVGVVTKPFNFEGKRRARSAESGI 157

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
             :|| |:||||||:|||||  :|: :||  :||: ||:|| :||:|:|||:| ||||||
Db        158 EELQKHVDTLIVIPNQNLFLIANPNTTFKQAFQMADEVLQQGVRGITDLMVCPGLINL 215


RESULT 3
Q9M7M6
ID   Q9M7M6       PRELIMINARY;      PRT;   401 AA.
AC   Q9M7M6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ OR MSFTSZ-MT.
OS   Mallomonas splendens.
OC   Eukaryota; stramenopiles; Chrysophyceae; Synurales; Mallomonas.
OX   NCBI_TaxID=52552;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MUCC 294;
RX   MEDLINE=20144160; PubMed=10678836;
RA   Beech P.L., Nheu T., Schultz T., Herbert S., Lithgow T., Gilson P.R.,
RA   McFadden G.I.;
RT   "Mitochondrial FtsZ in a Chromophyte Alga.";
RL   Science 287:1276-1279(2000).
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; AF120116; AAF35432.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
SQ   SEQUENCE   401 AA;  42337 MW;  E7B6A08C34A754BB CRC64;

  Query Match             59.5%;  Score 527;  DB 10;  Length 401;
  Best Local Similarity   58.4%;  Pred. No. 3.3e-39;
  Matches  104;  Conservative   27;  Mismatches   47;  Indels    0;  Gaps    0;

Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              | :| |||: |||| | |   |  :|: |||: |:||| |: || |:|:||: | :| |
Db        100 ARKLSGVEFVCANTDAQHLSTCLTENKLQLGKESTQGLGCGANPESGRRAAEESKEEIAR 159

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
               : |:||:||| |||||| |||||||| |  |  |||| ||: ||  || :| |||| |:
Db        160 YIADANMVFITAGMGGGTGTGAAPVVAEVCMEKDILTVAVVTKPFSFEGKHRARLANEGI 219

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
             : |  |||||::|||||:  | : ||:| :||  |||:|| ||| :|||:||||||||
Db        220 RSLEDRVDTLIIIPNQNIFKLINASTSMADAFGLADDILLAGVKSITDLMVRPGLINL 277


RESULT 4
Q9AQ38
ID   Q9AQ38       PRELIMINARY;      PRT;   421 AA.
AC   Q9AQ38;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Ehrlichia chaffeensis.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Anaplasmataceae; Ehrlichia.
OX   NCBI_TaxID=945;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lee K.N., Massung R.F., Padmalayam I., Baumstark B.;
RT   "Characterization of the ftsZ gene in Ehrlichia chaffeensis, the HGE
RT   agent and Rickettsia rickettsii.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; AF221944; AAK00615.2; -.
```

```
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
SQ   SEQUENCE   421 AA;  45664 MW;  E73CE4B1536255FA CRC64;

  Query Match             59.3%;  Score 525;  DB 2;  Length 421;
  Best Local Similarity   60.5%;  Pred. No. 5.3e-39;
  Matches  107;  Conservative   26;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | | || |:|||| ||| ||:  || ||  :|||||||| ||:|: :||:   :|
Db         36 SNLHGVNFVVANTDAQALELSLSEKKIQLGIGLTKGLGAGSLPEVGRGAAEESINEIIEE 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
              : |||||||| |||||| ||||||:| ||:|  |||:|||: ||  || :| | |  |::
Db         96 ISDSNMLFITAGMGGGTGTGAAPVIARVAKENKILTIGVVTKPFHFEGAHRMRTAEFGLE 155

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || :|||||||:|||||  :|:: ||  :||: || ||  ||:|:|||:| ||||||
Db        156 ELQRYVDTLIVIPNQNLFRIANEKTTFADAFKLADTVLHTGVRGITDLMVMPGLINL 212


RESULT 5
Q9AQ36
ID   Q9AQ36      PRELIMINARY;      PRT;   452 AA.
AC   Q9AQ36;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lee K.N., Massung R.F., Padmalayam I., Baumstark B.;
RT   "Characterization of the ftsZ gene in Ehrlichia chaffeensis, the HGE
RT   agent and Rickettsia rickettsii.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; AF221946; AAK00617.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
SQ   SEQUENCE   452 AA;  48365 MW;  5DCE41568630EBE0 CRC64;

  Query Match             59.2%;  Score 524;  DB 2;  Length 452;
  Best Local Similarity   56.2%;  Pred. No. 7.1e-39;
  Matches  100;  Conservative   38;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          1 ASQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQR 60
              :: |:|  |:||||| |:|  ||  :|| ||   |:|||||: ||:|  :|:: : :|:
Db         35 SANLQGANFVVANTDAQSLEHSLCTNKIQLGVSTTRGLGAGASPEVGALAAQESESEIRN 94

Qy         61 MLQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGV 120
              |::|||:||| |||||| ||:|||:| :|:|||||||||||: ||  || :| : |: |:
Db         95 YLENSNMVFITAGMGGGTGTGSAPVIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGL 154

Qy        121 KELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || ::|||||:|||||  :|:: ||  :||: ||||  ||:||||:: ||||||
Db        155 IELQQFVDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINL 212


RESULT 6
Q9AQ37
ID   Q9AQ37      PRELIMINARY;      PRT;   400 AA.
AC   Q9AQ37;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Anaplasmataceae; Anaplasma.
OX   NCBI_TaxID=948;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lee K.N., Massung R.F., Padmalayam I., Baumstark B.;
RT   "Characterization of the ftsZ gene in Ehrlichia chaffeensis, the HGE
RT   agent and Rickettsia rickettsii.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; AF221945; AAK00616.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
SQ   SEQUENCE   400 AA;  42001 MW;  CDE8730E876FDC9B CRC64;

  Query Match             58.9%;  Score 521;  DB 2;  Length 400;
  Best Local Similarity   59.3%;  Pred. No. 1.1e-38;
  Matches  105;  Conservative   30;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:||||| |||  ||:  || || ::|||||||| ||:|: :||:   :|
Db         36 SCLQGVNFVVANTDAQALDCSLSEKKIQLGMNLTKGLGAGSLPEIGRGAAEESIEEIIAE 95

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVK 121
              : |||||||| |||||| |||||||:|  |::  |||||||: ||  || :| : |  |::
Db         96 ISDSNMLFITAGMGGGTGTGAAPVIAKAAKDSKILTVGVVTRPFHFEGAHRMKTAEYGLE 155

Qy        122 ELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              || |:|||||:|||||  :|:::||  :||: || ||  ||:|:|||:| ||||||
Db        156 ELQKHVDTLIVIPNQNLFRIANENTTFADAFKLADTVLHTGVRGITDLMVMPGLINL 212


RESULT 7
Q8R9H2
ID   Q8R9H2      PRELIMINARY;      PRT;   357 AA.
AC   Q8R9H2;
```

```
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division GTPase.
GN   FTSZ OR TTE1639.
OS   Thermoanaerobacter tengcongensis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
OX   NCBI_TaxID=119072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MB4T / JCM11007;
RX   MEDLINE=21992816; PubMed=11997336;
RA   Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA   Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA   Tan H., Chen R., Wang J., Yu J., Yang H.;
RT   "A complete sequence of T. tengcongensis genome.";
RL   Genome Res. 12:689-700(2002).
DR   EMBL; AE013119; AAM24841.1; -.
KW   Complete proteome.
SQ   SEQUENCE   357 AA;  37777 MW;  D04B070F42493321 CRC64;

  Query Match           58.8%;  Score 520;  DB 16;  Length 357;
  Best Local Similarity  56.6%;  Pred. No. 1.2e-38;
  Matches  99;  Conservative  36;  Mismatches  40;  Indels  0;  Gaps  0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  ||| |||  | |  || :|: :|||||||: ||:||::||: : :|:| ::
Db         35 VKGVEFIAINTDKQALSLSKAETKIQIGEKLTKGLGAGANPEIGKKAAEESREEIERAIK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:||| |||||| ||||||| :|:|||||||||: ||  ||  |   |  |:::|
Db         95 GADMIFITAGMGGGTGTGAAPVVAEIAKELGILTVGVVTKPFTFEGRKRMAQAEMGIEDL 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              |||| || :||  || : :| |:||:||: ||||| :||:|::||  |||:|:
Db        155 KKYVDALITIPNDRLLQVVEKKTSMLDAFKLADDVLRQGVQGISDLIAVPGLVNV 209


RESULT 8
O85474
ID   O85474       PRELIMINARY;      PRT;   370 AA.
AC   O85474;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Clostridium lentocellum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=29360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC49066;
RX   MEDLINE=98374332; PubMed=9707627;
RA   Angert E.R., Losick R.M.;
RT   "Propagation by sporulation in the guinea pig symbiont Metabacterium
RT   polyspora.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:10218-10223(1998).
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; AF067822; AAC32265.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
SQ   SEQUENCE   370 AA;  39449 MW;  E7E81B612A992586 CRC64;

  Query Match           58.8%;  Score 520;  DB 2;  Length 370;
  Best Local Similarity  58.9%;  Pred. No. 1.3e-38;
  Matches  103;  Conservative  30;  Mismatches  42;  Indels  0;  Gaps  0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |||||||  ||| ||| || || || :|: :|:|||||: ||:| :|||: : :|   ::
Db         35 LEGVEFITVNTDHQALARSGAPAKIQIGEKMTRGLGAGANPEIGTKSAEESREEILTAIK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|||| |||||| ||||||:||:|:| ||||||||: ||  ||  |   |  |: ||
Db         95 GADMLFITAGMGGGTGTGAAPVIASIAKEEGILTVGVVTKPFSFEGRKRMINAEKGIAEL 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||:|:||  :| : || |||::||  ||||| :||:|:|||| ||:|||
Db        155 KQNVDTLVVIPNDKILQVIDKKTTMVDAFSKADDVLQQGVQGITDLISNPGIINL 209


RESULT 9
P77997
ID   P77997       PRELIMINARY;      PRT;   231 AA.
AC   P77997;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Cell division protein ftsZ (Fragment).
GN   FTSZ.
OS   Wolbachia sp.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Wolbachieae; Wolbachia.
OX   NCBI_TaxID=956;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96392953; PubMed=8799741;
RA   Tsagkarakou A., Guillemaud T., Rousset F., Navajas M.;
RT   "Molecular identification of a Wolbachia endosymbiont in a Tetranychus
RT   urticae strain (Acari: Tetranychidae).";
RL   Insect Mol. Biol. 5:217-221(1996).
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; U37260; AAC44314.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 2.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
FT   NON_TER       1      1
FT   NON_TER     231    231
SQ   SEQUENCE   231 AA;  24428 MW;  3BCB3F6C9DE936E0 CRC64;

  Query Match           58.4%;  Score 517;  DB 2;  Length 231;
  Best Local Similarity  56.1%;  Pred. No. 1.3e-38;
  Matches  106;  Conservative  31;  Mismatches  40;  Indels  12;  Gaps  1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:||||| ||| :||   || || ::|||||||: |::|| :||:   :|
```

```
Db          8 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDVGKGAAEESIDEIMEH 67

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
              ::||:||||| |||||| ||||||:|  |||                |||||||: ||  ||
Db         68 IKDSHMLFITAGMGGGTGTGAAPVIAKAAREARAVVKDKGAKEKKILTVGVVTKPFGFEG 127

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                | |:|  |::|| |||||||:|||||  :|:: || ::||: ||:||  |::|||||
Db        128 VRRMRIAELGLEELQKYVDTLIVIPNQNLFRIANEKTTXVDAFQLADNVLHIGIRGVTDL 187

Qy        170 IVRPGLINL 178
              :: ||||||
Db        188 MIMPGLINL 196


RESULT 10
O08390
ID   O08390      PRELIMINARY;      PRT;   289 AA.
AC   O08390;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ (Fragment).
GN   FTSZ.
OS   Wolbachia sp. 123B.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Wolbachieae; Wolbachia.
OX   NCBI_TaxID=59580;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=123B;
RA   Jager C.R., Pintureau B., Heddi A.;
RT   "Comparison between phylogenetic trees of some Trichogramma species
RT   and their Wolbachia endosymbionts.";
RL   Russ. Entomol. J. 7:163-168(1998).
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; U95751; AAB54069.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
FT   NON_TER       1      1
FT   NON_TER     289    289
SQ   SEQUENCE   289 AA;  30685 MW;  688372B39762D763 CRC64;

  Query Match             58.4%;  Score 517;  DB 2;  Length 289;
  Best Local Similarity   56.6%;  Pred. No. 1.7e-38;
  Matches  107;  Conservative   29;  Mismatches   41;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:||||| ||| :||   || || ::|||||||: |::|| :||:   :|
Db          2 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDVGKGAAEESIDEIMEH 61

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
              ::||:||||| |||||| ||||||:|  |||              ||||||: ||  ||
Db         62 IKDSHMLFITAGMGGGTGTGAAPVIAKAAREARAAVKDRAPKEKKILTVGVVTKPFGFEG 121

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                | |:|  |::|| |||||||:|||||  :|:: ||  :||: ||:||  |::|||||
Db        122 VRRMRIAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFSDAFKLADNVLHIGIRGVTDL 181

Qy        170 IVRPGLINL 178
              :| ||||||
Db        182 MVMPGLINL 190


RESULT 11
O08392
ID   O08392      PRELIMINARY;      PRT;   289 AA.
AC   O08392;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ (Fragment).
GN   FTSZ.
OS   Wolbachia sp. t191.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Wolbachieae; Wolbachia.
OX   NCBI_TaxID=59582;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=T191;
RA   Jager C.R., Pintureau B., Heddi A.;
RT   "Comparison between phylogenetic trees of some Trichogramma species
RT   and their Wolbachia endosymbionts.";
RL   Russ. Entomol. J. 7:163-168(1998).
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; U95753; AAB54071.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
FT   NON_TER       1      1
FT   NON_TER     289    289
SQ   SEQUENCE   289 AA;  30687 MW;  1640F8E382978512 CRC64;

  Query Match             58.2%;  Score 515;  DB 2;  Length 289;
  Best Local Similarity   56.6%;  Pred. No. 2.6e-38;
  Matches  107;  Conservative   28;  Mismatches   42;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |||||| ||| :||   || || ::|||||||: |::|| :||:   :|
Db          2 SNLQGVNFIVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDIGKGAAEESIDEIMEH 61

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
              ::||:||||| |||||| ||||||:|  |||              ||||||: ||  ||
Db         62 IRDSHMLFITAGMGGGTGTGAAPVIAKAAREARAVVKDKGAKEKKILTVGVVTKPFGFEG 121

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                | | |  |::|| |||||||:|||||  :|:: ||  :||: ||:||  |::|||||
Db        122 VRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFADAFQLADNVLHIGIRGVTDL 181

Qy        170 IVRPGLINL 178
              :: ||||||
Db        182 MIMPGLINL 190


RESULT 12
O08471
ID   O08471      PRELIMINARY;      PRT;   289 AA.
```

```
AC   O08471;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ (Fragment).
GN   FTSZ.
OS   Wolbachia sp.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Wolbachieae; Wolbachia.
OX   NCBI_TaxID=956;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M36, AND 1148;
RA   Jager C.R., Pintureau B., Heddi A.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; U95754; AAB54072.1; -.
DR   EMBL; U95750; AAB54068.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
FT   NON_TER       1      1
FT   NON_TER     289    289
SQ   SEQUENCE   289 AA;  30701 MW;  BF0D6BEECB430A07 CRC64;

  Query Match             58.1%;  Score 514;  DB 2;  Length 289;
  Best Local Similarity   56.1%;  Pred. No. 3.2e-38;
  Matches  106;  Conservative   29;  Mismatches   42;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:|||| ||| :||   || || ::|||||||: |::|| :||:   :|
Db          2 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDIGKGAAEESIDEIMEH 61

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
              ::||:|||| |||||| ||||||:|  |||                ||||||: ||  ||
Db         62 IRDSHMLFITAGMGGGTGTGAAPVIAKAAREARAVVKDKGAKEKKILTVGVVTKPFGFEG 121

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                | | |  |::|| ||||||||:|||||  :|:: ||  :||: ||:||  |::|||||
Db        122 VRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFADAFQLADNVLHIGIRGVTDL 181

Qy        170 IVRPGLINL 178
              :: ||||||
Db        182 MIMPGLINL 190


RESULT 13
O08389
ID   O08389      PRELIMINARY;      PRT;   289 AA.
AC   O08389;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ (Fragment).
GN   FTSZ.
OS   Wolbachia sp. 1032.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Wolbachieae; Wolbachia.
OX   NCBI_TaxID=59578;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1032;
RA   Jager C.R., Pintureau B., Heddi A.;
RT   "Comparison between phylogenetic trees of some Trichogramma species
RT   and their Wolbachia endosymbionts.";
RL   Russ. Entomol. J. 7:163-168(1998).
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; U95749; AAB54067.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
FT   NON_TER       1      1
FT   NON_TER     289    289
SQ   SEQUENCE   289 AA;  30702 MW;  B50A70FF77E41111 CRC64;

  Query Match             58.1%;  Score 514;  DB 2;  Length 289;
  Best Local Similarity   56.1%;  Pred. No. 3.2e-38;
  Matches  106;  Conservative   29;  Mismatches   42;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:|||| ||| :||   || || ::|||||||: |::|| :||:   :|
Db          2 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDIGKGAAEESIDEIMEH 61

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
              ::||:|||| |||||| ||||||:|  |||                ||||||: ||  ||
Db         62 IRDSHMLFITAGMGGGTGTGAAPVIAEAAREARAVVKDKGAKEKKILTVGVVTKPFGFEG 121

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
                | | |  |::|| ||||||||:|||||  :|:: ||  :||: ||:||  |::|||||
Db        122 VRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFADAFQLADNVLHIGIRGVTDL 181

Qy        170 IVRPGLINL 178
              :: ||||||
Db        182 MIMPGLINL 190


RESULT 14
O08391
ID   O08391      PRELIMINARY;      PRT;   289 AA.
AC   O08391;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division protein ftsZ (Fragment).
GN   FTSZ.
OS   Wolbachia sp. MB35.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Wolbachieae; Wolbachia.
OX   NCBI_TaxID=59581;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MB35;
RA   Jager C.R., Pintureau B., Heddi A.;
RT   "Comparison between phylogenetic trees of some Trichogramma species
RT   and their Wolbachia endosymbionts.";
RL   Russ. Entomol. J. 7:163-168(1998).
CC   -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
CC       ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
```

```
CC       THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
CC       AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
CC       BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
CC   -!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
CC       OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
DR   EMBL; U95752; AAB54070.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; GTP-binding; Septation.
FT   NON_TER       1      1
FT   NON_TER     289    289
SQ   SEQUENCE   289 AA;  30700 MW;  3D874344EB4DE409 CRC64;

  Query Match             58.1%;  Score 514;  DB 2;  Length 289;
  Best Local Similarity   56.1%;  Pred. No. 3.2e-38;
  Matches  106;  Conservative   29;  Mismatches   42;  Indels   12;  Gaps    1;

Qy          2 SQLEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRM 61
              | |:|| |:||||| ||| :||   || || ::|||||||: |:::|| :||: . :|
Db          2 SNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALPDIGKGAAEESIDEIMEH 61

Qy         62 LQDSNMLFITGGMGGGTCTGAAPVVASVAREL-----------GILTVGVVSTPFRSEG 109
              ::||:||||| |||||| ||||||:|  |||             |||||||: ||  ||
Db         62 IRDSHMLFITAGMGGGTGTGAAPVIAKAAREARAVVKDKGAKEKKILTVGVVTKPFGFEG 121

Qy        110 PNRTRLANAGVKELAKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDL 169
               | | |  |:::|| |||||||:|||||  :|:: ||  :||: ||:||  |::|||||
Db        122 VRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTFADAFQLADNVLHIGIRGVTDL 181

Qy        170 IVRPGLINL 178
              :: ||||||
Db        182 MIMPGLINL 190


RESULT 15
Q8Y5M5
ID   Q8Y5M5      PRELIMINARY;      PRT;   391 AA.
AC   Q8Y5M5;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   FtsZ protein.
GN   FTSZ OR LMO2032.
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGD-E / SEROVAR 1/2A;
RX   MEDLINE=21537279; PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL591982; CAD00110.1; -.
DR   ListiList; LMO02032; -.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   391 AA;  41350 MW;  6B9F75B2B1629C58 CRC64;

  Query Match             57.6%;  Score 510;  DB 16;  Length 391;
  Best Local Similarity   56.0%;  Pred. No. 1.1e-37;
  Matches   98;  Conservative   35;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              ::|||||  ||| |||  : |  |: :|  :|:|||||: ||:||::||: :  |:  |:
Db         35 VQGVEFISVNTDAQALNLAKAETKLQIGTKLTRGLGAGAVPEIGKKAAEESREQIEEALK 94

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                |:|:|:| |||||| ||||||:| :|:|:|: ||||||: ||  ||| ||: |  | : :
Db         95 GSDMVFVTAGMGGGTGTGAAPVIAQIAKEMGALTVGVVTRPFGFEGPKRTKQALTGTEAM 154

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
               : ||||||:||  || : ||:| |||||| ||:|| :||:|::||| ||||||
Db        155 KEAVDTLIVIPNDRLLQIVDKNTPMLEAFREADNVLRQGVQGISDLIAVPGLINL 209


Search completed: June  2, 2003, 08:41:12
Job time : 85 secs
```

```
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding.
FT   NP_BIND     101    109       GTP (POTENTIAL).
SQ   SEQUENCE   407 AA;  41809 MW;  01D2898E013220C5 CRC64;

  Query Match             54.8%;  Score 485;  DB 1;  Length 407;
  Best Local Similarity   53.1%;  Pred. No. 3.4e-34;
  Matches   93;  Conservative   38;  Mismatches   44;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||||  ||| |||  | |  |: :|:::|:|||||: | :|::::||  : :|: :|:
Db         32 LKGVEFIAINTDAQALLMSDADVKLDVGRELTRGLGAGANPAVGRKAAEDHREEIEEVLK 91

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::|:|:| | |||| || ||||::|| || ||:|||: ||  ||  |   |  |: ||
Db         92 GADMVFVTAGEGGGTGTGRAPVVANIARSLGALTIGVVTRPFTFEGRRRANQAEDGIAEL 151

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              : ||||||:||  ||:::|:  ::|:||: || ||| ||:|:|||| ||||||
Db        152 REEVDTLIVIPNDRLLSISDRQVSVLDAFKSADQVLLSGVQGITDLITTPGLINL 206


RESULT 15
FTSZ_AGRTU
ID   FTSZ_AGRTU     STANDARD;      PRT;   583 AA.
AC   O30992;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Cell division protein ftsZ.
GN   FTSZ.
OS   Agrobacterium tumefaciens.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
OX   NCBI_TaxID=358;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A136;
RX   MEDLINE=98012980; PubMed=9352931;
RA   Ma X., Sun Q., Wang R., Singh G., Jonietz E.L., Margolin W.;
RT   "Interactions between heterologous FtsA and FtsZ proteins at the FtsZ
RT   ring.";
RL   J. Bacteriol. 179:6788-6797(1997).
CC   -!- FUNCTION: This protein is essential to the cell-division process.
CC       It seems to assemble into a dynamic ring on the inner surface of
CC       the cytoplasmic membrane at the place where division will occur,
CC       and the formation of the ring is the signal for septation to
CC       begin. Binds to and hydrolyzes GTP (By similarity).
CC   -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Assembles at the inner surface
CC       of the cytoplasmic membrane (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF024659; AAC45821.1; -.
DR   HSSP; Q57816; 1FSZ.
DR   InterPro; IPR000158; FtsZ.
DR   InterPro; IPR003008; Tubulin_FtsZ.
DR   Pfam; PF00091; tubulin; 1.
DR   PRINTS; PR00423; CELLDVISFTSZ.
DR   TIGRFAMs; TIGR00065; ftsZ; 1.
DR   PROSITE; PS01134; FTSZ_1; 1.
DR   PROSITE; PS01135; FTSZ_2; 1.
KW   Cell division; Septation; GTP-binding.
FT   NP_BIND     107    115       GTP (POTENTIAL).
SQ   SEQUENCE   583 AA;  61936 MW;  3F415EB5FBCAA1CB CRC64;

  Query Match             54.5%;  Score 482;  DB 1;  Length 583;
  Best Local Similarity   57.1%;  Pred. No. 9e-34;
  Matches  100;  Conservative   26;  Mismatches   49;  Indels    0;  Gaps    0;

Qy          4 LEGVEFIVANTDCQALGRSLAPHKITLGKDITKGLGAGSKPELGKRSAEQQKVDIQRMLQ 63
              |:|||:|:||||| |||  : |   | || ::|:|||||||:||:|: :|||:   :|   |
Db         38 LQGVDFVVANTDAQALTMTKADRVIQLGVNVTEGLGAGSQPEVGRAAAEECIDEIIDHLN 97

Qy         64 DSNMLFITGGMGGGTCTGAAPVVASVARELGILTVGVVSTPFRSEGPNRTRLANAGVKEL 123
                ::| |:| |||||| |||||||  ||  ||||||||: ||  ||  | |||  |::||
Db         98 GTHMCFVTAGMGGGTGTGAAPVVAQAARNKGILTVGVVTKPFHFEGGRRMRLAEQGIEEL 157

Qy        124 AKYVDTLIVVPNQNLLALADKSTTMLEAFRYADDVLLEGVKGVTDLIVRPGLINL 178
              | ||||||:|||||  :|:  ||  :||  || ||  ||  :|||:|: |||||
Db        158 QKSVDTLIVIPNQNLFRIANDKTTFADAFAMADQVLYSGVACITDLMVKEGLINL 212


Search completed: June  2, 2003, 08:39:32
Job time : 28 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:47:19 ; Search time 2357 Seconds
(without alignments)
2197.836 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09770509/runat_27052003_083648_18820/app_query.fasta_1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09770509_@CGN_1_1_2496_@runat_27052003_083648_18820 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	547	61.8	3824	1	AF492457	AF492457 Azospiril
2	527	59.5	1404	8	AF120116	AF120116 Mallomona
c 3	527	59.5	38510	1	AF179611	AF179611 Zymomonas
4	525	59.3	1542	1	AF221944	AF221944 Ehrlichia
5	524	59.2	1540	1	AF221946	AF221946 Rickettsi
6	524	59.2	8328	1	AE008653	AE008653 Rickettsi
7	521	58.9	1474	1	AF221945	AF221945 Ehrlichia
8	520	58.8	1202	1	AF067822	AF067822 Clostridi
c 9	520	58.8	11683	1	AE013119	AE013119 Thermoana
10	518	58.5	279110	1	RPXX03	AJ235272 Rickettsi
11	517	58.4	697	1	WSU37260	U37260 Wolbachia s
12	517	58.4	868	1	WSU95751	U95751 Wolbachia s
13	515	58.2	868	1	WSU95753	U95753 Wolbachia s
14	514	58.1	868	1	WSU95749	U95749 Wolbachia s
15	514	58.1	868	1	WSU95750	U95750 Wolbachia s
16	514	58.1	868	1	WSU95752	U95752 Wolbachia s
17	514	58.1	868	1	WSU95754	U95754 Wolbachia s
c 18	511	57.7	303249	1	AP001515	AP001515 Bacillus
19	510	57.6	1194	1	DMFTSZ	X71906 Wolbachia s
20	510	57.6	213251	6	AX413015	AX413015 Sequence
c 21	510	57.6	258650	1	AL596171	AL596171 Listeria
c 22	510	57.6	295050	1	AL591982	AL591982 Listeria
c 23	510	57.6	349980	6	AX417047	AX417047 Sequence
24	509	57.5	950	6	AX436249	AX436249 Sequence
25	507	57.3	13165	1	AE007679	AE007679 Clostridi
26	506	57.2	8834	1	BBCDG	X96685 B.burgdorfe
27	506	57.2	15272	1	BORFTSA	L76303 Borrelia bu
c 28	506	57.2	34817	1	BBU43739	U43739 Borrelia bu
c 29	506	57.2	44380	1	AE001137	AE001137 Borrelia
30	505	57.1	1877	1	CCU40273	U40273 Caulobacter
c 31	505	57.1	10487	1	AE005922	AE005922 Caulobact
32	503	56.8	1864	1	AF061746	AF061746 Bartonell
33	503	56.8	1864	6	AX191760	AX191760 Sequence
34	502	56.7	1134	6	AX431765	AX431765 Sequence
35	499	56.4	1893	1	AF061747	AF061747 Bartonell
36	498	56.3	4022	1	BACFTSZA	M22630 B.subtilis
37	498	56.3	208230	1	BSUB0008	Z99111 Bacillus su
38	497	56.2	1052	1	AF282845	AF282845 Wolbachia
39	496	56.0	1859	1	AF141018	AF141018 Bartonell
c 40	495	55.9	296750	1	AP003191	AP003191 Clostridi
41	494	55.8	1782	1	AF007266	AF007266 Bartonell
42	493	55.7	1201	1	AF067823	AF067823 Clostridi
c 43	493	55.7	13423	1	AE006416	AE006416 Lactococc
44	492	55.6	11551	1	EHY13922	Y13922 Enterococcu
45	491	55.5	1055	1	AF220605	AF220605 Wolbachia

ALIGNMENTS

RESULT 1

```
AF492457
LOCUS       AF492457               3824 bp    DNA     linear   BCT 11-APR-2002
DEFINITION  Azospirillum brasilense D-alanine:D-alanine ligase (ddlB), partial
            cds; and cell division fts gene cluster, partial sequence.
ACCESSION   AF492457
VERSION     AF492457.1  GI:20136382
KEYWORDS    .
SOURCE      Azospirillum brasilense.
  ORGANISM  Azospirillum brasilense
            Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
            Azospirillum.
REFERENCE   1  (bases 1 to 3824)
  AUTHORS   Jofre,E.C., Lagares,A. and Mori,G.B.
  TITLE     A ddlB (D-alanine-D-alanine ligase) Tn5 mutant of Azospirillum
            brasilense is altered in exopolysaccharide production, salt
            tolerance, and maize root colonization
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3824)
  AUTHORS   Jofre,E.C., Lagares,A. and Mori,G.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2002) Natural Sciences, Universidad Nacional de
            Rio Cuarto, Route 36 KM 601, Rio Cuarto, Cordoba 5800, Argentina
FEATURES             Location/Qualifiers
     source          1. .3824
                     /organism="Azospirillum brasilense"
                     /strain="Cd"
                     /db_xref="taxon:192"
     gene            <1. .632
                     /gene="ddlB"
     CDS             <1. .632
                     /gene="ddlB"
                     /codon_start=3
                     /transl_table=11
                     /product="D-alanine:D-alanine ligase"
                     /protein_id="AAM11649.1"
                     /db_xref="GI:20136383"
                     /translation="KALLAPVGVRSPKGLVLTKGELTGGAHPMPAPYIVKPVDEGSTV
                     GVTLVREGQNSPVGDAWTFGERALVEEFIPGRELTVGVMGGLDGECRPLTVTEIRFEA
                     QVYDYTAKYSAGHAVHTVPAQIPEHVAEEAKRLAVLAHRTLGCRGVSRSDFRWDDRKS
                     GTDGLFFLEINNQPGMTPLSLVPEQAASVGLTYGALVAWMVENAACQLD"
     gene            617. .1579
                     /gene="ftsQ"
     CDS             617. .1579
                     /gene="ftsQ"
                     /note="may be involved in septum formation"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein FtsQ"
                     /protein_id="AAM11650.1"
                     /db_xref="GI:20136384"
                     /translation="MSTRLSIDPQFRPGFGGSANRGRDDRPVPPRAMAKSAQKGNRRR
                     AWPRWTRPAVKAAILLTPVLVVAGMAASAWQRGTIAETTAALQESLIQTSASAGFAIA
                     DVLVEGRTETDPASILRVLGVQRGDPILAVTLSDAKEKLESLPXVASXSIERHLPDIL
                     FVRLTERQPMAIWQHDRKFTVIDREGRPLADATELARRGNRRIETLPQVVGANAPMQV
                     PKLLAALDNVPALREKVSAASWVGDRRWDLKLKNGVVVKLPEARMQSALRQLAEMDAT
                     GQVLDRDIVAIDLRQNDRAVLQTSATAVLPWTEEENKKKPGKKT"
     gene            1662. .2924
                     /gene="ftsA"
     CDS             1662. .2924
                     /gene="ftsA"
                     /note="may be involved in anomalous filament growth and a
                     component of the septum; may interact with FtsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein FtsA"
                     /protein_id="AAM11651.1"
                     /db_xref="GI:20136385"
                     /translation="MGFNGAKKPKRATRGGIIAALDVGSTKVCCLIARVEDAGGVRIV
                     GIGHQIATGVRAGAIIDMEAAETSIGAAVHAAEQMANETIRDVIVNVSSPISHGFNAE
                     VPVSGQEVTESDVRRALAHARSLQVGPDQALVHAIPVGFALDGSRGIRDPKGMYGERL
                     GVQAHVITSPAGAVRNLQTCVARCHLDIEGLVASPYASGLACLVDDEMEMGSACIDMG
                     GGTTTISVFSEGTLVWSDCIRLGGNHVTNDIARGLTTPVVHAERMKTLHGSAINSPAD
                     EREMIDVPQVGEEERLGANNLPKSYLVRIIQPRLEEIFEHVRSRLEHSGYAKLVGRRV
                     VLTGGASQLPGMRELAQLILDKQVRIGRPTRITGLNEAQGGPSYSTAAGLLLHAVRNP
                     AELPVAGHEAGAGTGFFGRVGLWLRENL"
     gene            3067. .>3824
                     /gene="ftsZ"
     CDS             3067. .>3824
                     /gene="ftsZ"
                     /note="essential for the cell-division process; assembles
                     into a dynamic ring on inner surface of cytoplasmic
                     membrane at location of cell division; binds to and
                     hydrolyzes GTP"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein FtsZ"
                     /protein_id="AAM11652.1"
                     /db_xref="GI:20136386"
                     /translation="MINVTIPSIEPELKPRITVFGVGGAGGNAVNNMIKSNLEGVDFV
                     VGNTDAQALKGSLCEKRVQLGTTMTRGLGAGSKPDVGRASAEEQLEEIIGHLEGANMV
                     FITAGMGGGTGTGAAPVIARAARERGLLTVGVVTKPFHFEGAHRMRLAESGIAELQQY
                     VDTLIIIPNQNLFRIANEKTTFADAFKMADDVLHSGVRGVTDLMVMPGLINLDFADIR
                     SVMTEMGKAMMGTGEAGGERRAIEGRRGPHLQPAA"
BASE COUNT      664 a   1353 c   1260 g    544 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:                9.42e-41      Length:          3824
Score:                    547.00        Matches:         107
Percent Similarity:       77.40%        Conservative:    30
Best Local Similarity:    60.45%        Mismatches:      40
Query Match:              61.81%        Indels:          0
DB:                       1             Gaps:            0

US-09-770-509-2 (1-178) x AF492457 (1-3824)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||||||||||||:::|||:::|||   |||||||||   |||||||||
Db       3172 TCCAACCTCGAAGGCGTGGACTTCGTGGTCGGCAACACCGACGCGCAGGCCCTGAAGGGC 3231

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              ||||||          ::::::   ||||||      :::|||:::|||||||||||||||||
Db       3232 TCGCTCTGCGAGAAGCGCGTCCAGCTCGGCACCACCATGACGCGCGGCCTGGGCGCCGGC 3291

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              |||||||||:::::|||:::   |||||||||:::|||      :::|||
Db       3292 TCCAAGCCGGACGTCGGCCGGGCCTCCGCCGAGGAGCAGCTCGAGGAGATCATCGGTCAC 3351

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::   :::||||||:::|||||||||   ||||||||||||||||   ||||||
Db       3352 CTCGAAGGCGCCAACATGGTGTTCATCACCGCCGGCATGGGCGGCGGCACCGGCACGGGT 3411

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||:::|||      |||||||||   |||:::|||||||||||||||||||
Db       3412 GCGGCCCCGGTCATCGCCCGCGCCGCCCGCGAGCGCGGCCTGCTGACCGTCGGCGTGGTG 3471

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||   :::|||   |||||||||   :::|||:::
Db       3472 ACCAAGCCCTTCCACTTCGAGGGCGCGCACCGCATGCGGCTGGCCGAATCCGGCATCGCC 3531

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              ||||||   :::|||||||||||||||||||:::::|||||||||||||||      :::
Db       3532 GAGCTGCAGCAGTATGTTGACACGCTGATCATCATCCCGAACCAGAACCTGTTCCGCATC 3591

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              |||:::::::   ||||||      :::|||||||:::   ||||||||||||||
Db       3592 GCCAACGAGAAGACGACCTTCGCGGACGCCTTCAAGATGGCCGACGACGTTCTGCATTCC 3651

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||||||||||||||:::|||   ||||||||||||||||||
Db       3652 GGCGTGCGCGGCGTGACCGACCTGATGGTGATGCCCGGCCTCATCAACCTG 3702

RESULT 2
AF120116
LOCUS       AF120116               1404 bp    mRNA    linear   PLN 25-FEB-2000
DEFINITION  Mallomonas splendens FtsZ (MsFtsZ-mt) mRNA, nuclear gene encoding
```

```
          mitochondrial protein, complete cds.
ACCESSION   AF120116
VERSION     AF120116.1  GI:7024509
KEYWORDS    .
SOURCE      Mallomonas splendens.
  ORGANISM  Mallomonas splendens
            Eukaryota; stramenopiles; Chrysophyceae; Synurales; Mallomonas.
REFERENCE   1  (bases 1 to 1404)
  AUTHORS   Beech,P.L., Nheu,T., Schultz,T., Herbert,S., Lithgow,T.,
            Gilson,P.R. and McFadden,G.I.
  TITLE     Mitochondrial FtsZ in a chromophyte alga
  JOURNAL   Science 287 (5456), 1276-1279 (2000)
  MEDLINE   20144160
   PUBMED   10678836
REFERENCE   2  (bases 1 to 1404)
  AUTHORS   Beech,P.L., Schultz,T., Vuong,T., Gilson,P.R. and McFadden,G.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-1999) School of Botany, University of Melbourne,
            Grattan St, Parkville, Vic. 3052, Australia
FEATURES             Location/Qualifiers
     source          1. .1404
                     /organism="Mallomonas splendens"
                     /strain="MUCC 294"
                     /db_xref="taxon:52552"
     gene            1. .1404
                     /gene="MsFtsZ-mt"
     CDS             10. .1215
                     /gene="MsFtsZ-mt"
                     /note="putative mitochondrial division protein"
                     /codon_start=1
                     /product="FtsZ"
                     /protein_id="AAF35432.1"
                     /db_xref="GI:7024510"
                     /translation="MRITGANRILSLSRIRHFSDGASLNKAFLRSVKPGVKPEQYDSR
                     SGNSSQAQSTEHVKDKFVEPGNLRFRTGEYITEFLPKICVFGVGGGGCNAVNNMIARK
                     LSGVEFVCANTDAQHLSTCLTENKLQLGKESTQGLGCGANPESGRRAAEESKEEIARY
                     IADANMVFITAGMGGGTGTGAAPVVAEVCMEKDILTVAVVTKPFSFEGKHRARLANEG
                     IRSLEDRVDTLIIIPNQNIFKLINASTSMADAFGLADDILLAGVKSITDLMVRPGLIN
                     LDFADVRTVMSGMGHAIMGTGQAEGEDRAIRAANDALNNPLLGGDFSVRSAKGMLVNI
                     TGGKDLTLVEVDAAAQRITSEIEDEDANVIFGSSFDESLQGSIRVSIVATGIEAPGAA
                     AATAAPVIR"
BASE COUNT      404 a    278 c    373 g    349 t
ORIGIN

Alignment Scores:
Pred. No.:              2.05e-39        Length:         1404
Score:                  527.00          Matches:        104
Percent Similarity:     73.60%          Conservative:   27
Best Local Similarity:  58.43%          Mismatches:     47
Query Match:            59.55%          Indels:         0
DB:                     8               Gaps:           0

US-09-770-509-2 (1-178) x AF120116 (1-1404)

Qy          1 AlaSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
              |||   :::|||   |||||||||||||:::   ||||||||||||   |||   |||
Db        307 GCTAGGAAACTGTCTGGAGTAGAATTTGTCTGCGCGAATACTGACGCACAACATCTTTCA 366

Qy         21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
                 |||       :::|||:::   ||||||||||:::   |||:::|||||||||
Db        367 ACTTGCTTGACTGAAAACAAGCTTCAGCTAGGAAAAGAGTCAACCCAAGGATTGGGTTGC 426

Qy         41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
              |||:::   ||||||   |||:::|||:::|||||||:::   |||   :::|||   |||
Db        427 GGTGCAAACCCAGAATCAGGGCGTCGAGCCGCGGAAGAAAGTAAAGAGGAAATTGCCAGA 486

Qy         61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
                :::   |||:::|||||||:::|||||||||   ||||||||||||||||   |||
Db        487 TATATTGCAGATGCTAATATGGTATTTATAACTGCCGGTATGGGTGGTGGAACAGGAACA 546

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
              ||||||||||||||||||||   |||       |||       ||||||||||||   |||
Db        547 GGGGCCGCGCCTGTAGTAGCCGAGGTTTGCATGGAAAAGGACATTCTAACGGTGGCAGTG 606

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
              |||:::   ||||||      ||||||   :::|||   ||||||||||||   |||:::
Db        607 GTCACTAAACCATTTAGCTTTGAGGGGAAGCATCGCGCTCGCCTAGCAAACGAAGGAATA 666

Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
              :::   |||         ||||||||||||||:::::|||||||||||||:::
Db        667 AGGTCTCTCGAAGATCGTGTTGACACGCTAATAATAATTCCAAATCAAAATATATTCAAG 726

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
              |||   :::   ||||||:::|||   :::||||||       |||||||||:::||||||
Db        727 CTCATTAACGCGTCGACGTCAATGGCCGATGCGTTCGGCCTGGCTGACGACATTTTGTTG 786

Qy        161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                  |||||||||   :::||||||||||:::||||||||||||||||||||||||
Db        787 GCCGGCGTGAAGAGCATCACGGACCTGATGGTTCGGCCGGGACTGATCAACTTG 840

RESULT 3
AF179611/c
LOCUS       AF179611               38510 bp    DNA     linear   BCT 07-SEP-1999
DEFINITION  Zymomonas mobilis ZM4 fosmid clone 41E10, complete sequence.
ACCESSION   AF179611
VERSION     AF179611.1  GI:5834350
KEYWORDS    .
SOURCE      Zymomonas mobilis.
  ORGANISM  Zymomonas mobilis
            Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
            Zymomonas.
REFERENCE   1  (bases 1 to 38510)
  AUTHORS   Um,H.W. and Kang,H.S.
  TITLE     Sequence analysis of 41E10 fosmid clone of Zymomonas mobilis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 38510)
  AUTHORS   Um,H.W. and Kang,H.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-1999) Microbiology, Seoul National University,
            San 56-1 Shilimdong Kwanaku, Seoul 151-742, Korea
FEATURES             Location/Qualifiers
     source          1. .38510
                     /organism="Zymomonas mobilis"
                     /strain="ZM4"
                     /db_xref="taxon:542"
                     /clone="fosmid 41E10"
     gene            <1. .465
                     /gene="PPIASE"
     CDS             <1. .465
                     /gene="PPIASE"
                     /codon_start=1
                     /transl_table=11
                     /product="peptidyl prolyl cis trans isomerase"
                     /protein_id="AAD53917.1"
                     /db_xref="GI:5834351"
                     /translation="DPKGDGSGGSSLPNLPAEFNDMPHVRGEVSMARAQEKDSANSQF
                     FIVYQPTFSLDGEYTPIGRVVSGMEYVDAVEKGEPPAHPSIIVRASIGSDKIHPPTTE
                     EVKAAVEKQEAINLAEQQKQEEEVRRNAMRSFLPNHARLLPMQNRRRLRQSH"
     gene            502. .1530
                     /gene="queA"
     CDS             502. .1530
                     /gene="queA"
                     /note="QueA"
                     /codon_start=1
                     /transl_table=11
                     /product="S-adenosylmethionine tRNA ribosyltransferase"
                     /protein_id="AAD53918.1"
                     /db_xref="GI:5834352"
                     /translation="MRVDLFDFELPTESIALRPAVPRESARLLQVTSQEMDDYHVADL
                     PNLLQEGDLLIFNDTRVIPAYLEGKKNQARISANMIKREGLRQWQAFVRNAKRLKSDD
                     IVDFGQDVSARVVHRNDDGSVLFEFLGNEPVEILLERAGNMPLPPYIAHKRPADQQDR
                     QDYQTIFAEKEGAVAAPTASLHFTEELLQRLEEKGVKHTKVTLHVGAGTFLPMKVEDS
                     DDHVMHSEWGQISAETANAINETKKQGGRVIATGTTSLRLIDSAADETGLVHPFSDET
                     DIFITPGYRFRVIDGLMTNFHLPKSTLFMLVSALMGREKMLAAYQHAIDTGYRFYSYG
                     DSSLLLPK"
     gene            1596. .2174
```

/gene="fprA"
CDS 1596. .2174
/gene="fprA"
/note="PNP/PMP oxidase"
/codon_start=1
/transl_table=11
/product="pyridoxamine 5'-phosphate oxidase"
/protein_id="AAD53919.1"
/db_xref="GI:5834353"
/translation="MTQDPHKMFDEWMDEAKKAETDDPTAMALATASKQAFPCVRMML
LKGHTKDGFVFFTNLGSRKGHELLENPVATLLFHWKKLRRQVRIEGAATLISDEEADA
YFATRARKSQLGAWASEQSRPLPARDVFEKRIADIEARYEGKDVPRPPYWTGFRVSPI
RMEFWNDREFRLHERELFTLNDGRWQSEFLYP"
gene complement(2214. .3332)
/gene="ntrC"
CDS complement(2214. .3332)
/gene="ntrC"
/codon_start=1
/transl_table=11
/product="nitrogen assimilation regulatory protein"
/protein_id="AAD53920.1"
/db_xref="GI:5834354"
/translation="MASISSFRSASVVENTALIGESPAMREIRYLIQQVAPTSASVLI
TGPSGSGKEMVASSIHASSPRSQGRFIPVNCGAIPRDLLESELFGHEKGAFTGAVGQT
CGRFEEADGGTLFLDEIGDMPFDMQVKLLRVLEDNIITRVGGRQSIHVDTRIISATHR
DIHNAIEHKAFREDLFYRLAVFPIHLPSLADRSDDVPLLAEYFCDRLPENRKFCFSSD
SLHYLSQYSWPGNIRELRNLIERTSILYPDQKVGVEQTQKLLNLAVSSQPADRDALWN
LSSPSSEISSDRDLSLADDILEDDPDNESENLALFADFPIPGDAVKLKELVTELECRY
ITAALNSANNIVSDAARLLGLHRTTLIEKMNKYKISRS"
gene 4117. .5646
/gene="zm11orf1"
CDS 4117. .5646
/gene="zm11orf1"
/codon_start=1
/transl_table=11
/product="phosphatidylinositol glycan homolog"
/protein_id="AAD53921.1"
/db_xref="GI:5834355"
/translation="MANKDKSAASFSLKSIQASHLTVLLITLLALCLRMTAIFPIGYY
HPDEIFQYLEQAHRLAFGYSVIPWEYRYGMRGGLVPVLLAAAMKMGDIMNTQSFLYWV
LPHAITALLSLGVIWSAYKLGRLHSFFTGCLAMFVAATWYEIVLFAAHPLSESLSFAA
FMPAAYYLIAEEKKKSSLILSGFLMGLSALLRFQYIPSILFITLFTLRINVRQWGYFI
LGGLGAVFISSLSDIFIGEYPLLWLFKNIQQNLIFHRAEDYGTAPIYSYFFDIWTNWK
WFSLILPIPILVAAYRYPALFGTALVNLIFHMLVSHKEYRFIFLSSGIFIILAAIGSA
LLIEKATINFTPLHKKLAQIAAFTAWGACSFLCSTGNISHWYQYRGVLQAEMKARSLS
NLCGLTLYQIDYWESGGYSLLHRPIPIYISLLPSNMYSTRYPDPIALSPSLFSASNAII
SAASSGKNLPKEYRQVSCYPQKNGEKTEDICLFFRPSSCFKKETADNFEINKVLKNID
Q"
gene complement(5933. .7819)
/gene="zm11orf2"
CDS complement(5933. .7819)
/gene="zm11orf2"
/codon_start=1
/transl_table=11
/product="porin homolog"
/protein_id="AAD53922.1"
/db_xref="GI:5834356"
/translation="MKNNFKIAKVAQWVLALSIATTSAEAWADPLGNWLAKVNNHHKQ
KESASEQSVEQDDRTDTDNTVNTSPVANQPQAQTSQTVSQPSAVTPYQSVNNRESQAA
SQIPAADATDNDVRMSDNDNNDASGRSVDAPQMAEVSHQHVSPAPFPYASLTVGSSEL
QPETDPRLPVSKGGVVEENVENKGLVDSIVQSTRGRHNAFAWIGVGDNSVMLARDPKD
RGMFGTWLGIKPWLRDKGIDIQSSMGGVETAFNARGGDRHGLRTAYQWLVQASFDMER
LGVVKGGKFNFMMTKRGGGQIEQAINSTTLTEMQEVYGRGNIWRLIEGSYEQQIGRFN
IKVGRLTENNEFGMAMTGCSFMSLGFCGNQAGNLVGDYWYNWPVSNWAGRVTYQVSKE
VTIGAGVFDVNRKNIQGKVLHENFYMDNPFHPEAGLYIGEVKWAPKFGRENNLEGLFK
FDGWYTDDKSYNLRYYQDVAMQTGYSSQGPEVRGRYGAYVTAEQQVLRFDSHRQLTVA
FRALKNDGRTNKSTAEVSAYAILFGVSKERPDDWVGLGAGTTWLSNAFLKSQALSAQN
RHQEYYSELFYNIQLSPAVFLRPNIQYVVHPGGYYARGPVDTPYNYNSKNIILLGLKS
GWVF"
gene 8520. .10292
/gene="kefC"
CDS 8520. .10292
/gene="kefC"
/note="KefC; K+ efflux antiporter"
/codon_start=1
/transl_table=11
/product="glutathione-regulated potassium efflux system protein"
/protein_id="AAD53923.1"
/db_xref="GI:5834357"
/translation="MNSSLFSDSLVILGAAGLIIPAFARFRISPVIGFILAGAVLGPK
ALGSLVTEYPWLKHITISDPDGIEPFADFGIVLMLFSIGIELSFKRLWAMRQAVFGLG
TGKLLGCGFLIGAGLYALGENWAGAIGMGTGSQRIVNRFGAASSGNNGTCRKKPFAIL
LFEDLALVPIIFALVPWHLMGKIIAGTDLIATIWQGGATVIAILVLGRFALPTLFAQA
ARTKSAEVFLSISLLVVILASVATSAMGLSPIVGALTAGLLIAETDYHVEVEMVILPF
QRLALGVFLLTVGMSLDLRLIMANWQALALGVAGVVLVKTLVTSTLLRLSHVRLATAV
ESCLLMSAPSETTLVILATARQAQLIMPSTAAFWEIVTAIGLTITPHLAKISHKVSLK
LTLREHNLTLPSNEEPIEDIQNHTIVAGIGRVGKVIADVLTSEKKPYLAIDSDIDLVI
NAKRDGYHATFGDISNPVFIQQLGNKINAVVLAINDPVSMVNVVHRLHTIYPDLTIIA
RASDSRHAAQLFRAGATDIVAETLESSLRLAEATLISLGVSTGHAVAAIHDKRASMRE
NVLKAAEEGNMQAIRHTPSIMPEA"
gene complement(10368. .13034)
/gene="alaS"
CDS complement(10368. .13034)
/gene="alaS"
/note="AlaS"
/codon_start=1
/transl_table=11
/product="alanyl-tRNA synthetase"
/protein_id="AAD53924.1"
/db_xref="GI:5834358"
/translation="MITTNEIRRSFSRIFRKEWPSDSTIAPLVPQNDPTLMFVNAGMV
PFKNTFTGLESRPYKTATSSQKCVRAGGKHNDLDNVGYTARHHTFFEMLGNFSFGDYF
KERVIELAWGLITKEWGLDPERLCVTVYHTDEEAFNLWRKIAGLPEDRIIKIATSDNF
WSMGDTGPCGPCSEIFYDHGPEIPGASGLADEDGDRFVEIWNLVFMQYEQVNAETRLN
LPRPSIDTGMGLERIASVLQGVHDNYDTDTFKALIAASGELTHTATDGQFKASHRVIA
DHLRAAGFLVADGVLPANEGRGYVLRRIMRRAMRHAHLIGAKEPLMYRLVPALLSEMG
MAYPELVRAKALIEETLRLEETRFRQTLANGLKILKDETQHLKSGDTLPGAVAFRLYD
TYGFPYDLTADALRARNLTVDQAGFDAAMAEQRKAARAAWKGSGEKASDEIWFDIADQ
LGGTEFTGYTAEKGSGQIIALIKDGKRVETAKQGDDITIITNQTPFYGESGGQKGDIG
VITGNNDLKMTVTDTQKPLGRIHAHIAKIEKGEIKIGDDIQLQVDINHRNRLRANHSA
THLLHAALRDQLGQHVSQKGSMVSAERLRFDFSHQKALSDQELAAIEAEVNQQILNNS
VVTTRLMTPESAVEAGAMALFGEKYGNEVRVLSMGSCLNDQNEESSWSVELCGGTHVS
ALGQIGLFHIVSETAVSSGIRRIEAVTGEEARLWLVGRDRLLRETASILKAVPEEVPT
RTAAILDERRKCQRSLADQKALAFANANCGQGGNNAAPDKLGAYHFIGQVIEGLEPKA
LLGLIDENKKIIESGVIALITVNEGRASVAIGVSDSLKGKISAVDLVRKAVETLGGKG
GGGRPDMAQGGGPNGNEAAQALEAVKALLEKA"
gene complement(13374. .14303)
/gene="zm11orf3"
CDS complement(13374. .14303)
/gene="zm11orf3"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAD53925.1"
/db_xref="GI:5834359"
/translation="MDDKDQEHDAMNSHGRPQENNTYNRPFGRSNAHGSSRDDTSDFY
DHHSADADDGDEDDSLPWLEPVEDDDNGHGGNIFQIVIVALVALLTLALLGIGLYWWF
HRPPAVSGNASVITAEPGPYKTKPREPGGMKIEGEGESAYAASEGRDINSSIDTSVQP
EQPMTIAPKQAEQPITPPPPASAPVTKTRPDVVENEPEMSDMTGDTRHSTATHPAAPR
PAAPVKHEKEVKPAEHPVKAAEREKPPVAVVKPAAKVEKPEAEHQAEVAKPKSREDSV
AGVMSSLHNDAHPADEAKKAPAGCRCYPVGCLW"
gene complement(14307. .16049)
/gene="argS"
CDS complement(14307. .16049)
/gene="argS"
/note="ArgS"
/codon_start=1
/transl_table=11
/product="arginyl-tRNA synthetase"
/protein_id="AAD53926.1"
/db_xref="GI:5834360"
/translation="MDYFNRDPLSPFCRALSSILDTFEAEGLLPAGLDRSRVAVEPPR
DPSHGDLAINAAMVLAKAAKTAPRALAEKLVESLSQLEGVESVEIAGPGFINIRLTDD
CWRKELTHILEQGQDYGRSDIGQNLRVNVEYVSANPTGPMHMGHCRGAVVGDALASLL
DFAGYKVTREYYVNDAGQQVQTLARSAYMRYREALGEEITIPEGLYPGDYLVPVGKLL

```
Alignment Scores:
Pred. No.:                9.2e-38        Length:         38510
Score:                    527.00         Matches:        106
Percent Similarity:       75.84%         Conservative:   29
Best Local Similarity:    59.55%         Mismatches:     43
Query Match:              59.55%         Indels:         0
DB:                       1              Gaps:           0

US-09-770-509-2 (1-178) x AF179611 (1-38510)

Qy          1 AlaSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
              ||||||   ::::::||||||:::|||||||||||||||||||||   |||||||||
Db      22221 GCATCAGGTGTGCAGGGTGTCGATTTTATTGTAGCCAATACAGATGCCCAGGCCTTAAAT 22162

Qy         21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
                |||   |||      :::|||   ||||||         |||:::|||||||||||||
Db      22161 ATCTCTCCGGCAGAACAGCGGATTCAGCTTGGCCCGACGACGACGCAGGGCTTGGGGGCA 22102

Qy         41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
              ||||||:::|||||||:::||||||   :::|||||||:::            ||||||
Db      22101 GGATCGAGACCGGAAGTCGGGAAAGCGGCCGCAGAAGAAACGATCGAGCAGATTCAGGAA 22042

Qy         61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
              |||:::  :::  |||  ||||||      ||||||||||||||||||  |||
Db      22041 GCCTTGGAAGGCGCAAGAATGTGCTTTATCGCCGCTGGGATGGGGGGCGGCACCGGCACG 21982

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
              ||||||||||||||:::|||   ||||||||:::   ||||||||||||||||||||
Db      21981 GGTGCAGCTCCGGTCATCGCAAAAGTAGCGCGTGATCGTGGTATTCTGACAGTCGGTGTT 21922

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
              |||:::  ||||||      ||||||      |||  |||  |||  :::|||:::
Db      21921 GTGACCAAACCGTTCAACTTTGAAGGTAAAAGGCGGGCGCGGTCAGCTGAAAGCGGTATT 21862

Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
              :::||||||   |||:::||||||||||||||||||:::|||||||||||||||
Db      21861 GAAGAGCTGCAAAAGCATGTCGATACGCTGATTGTTATTCCAAACCAGAATTTGTTTTTG 21802

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
              :::|||:::  :::||||||      :::||||||:::  |||||||:::|||||||
Db      21801 ATTGCCAATCCGAACACGACCTTTAAGCAGGCCTTCCAGATGGCGGATGAGGTTTTGCAG 21742

Qy        161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::||||||:::|||:::||||||||||:::|||   ||||||||||||||||||
Db      21741 CAGGGTGTTCGCGGTATTACCGATTTGATGGTTTGCCCCGGTCTTATCAATCTT 21688

RESULT 4
AF221944
LOCUS       AF221944               1542 bp    DNA     linear   BCT 15-MAY-2001
DEFINITION  Ehrlichia chaffeensis cell division protein FtsZ (ftsZ) gene,
            complete cds.
ACCESSION   AF221944
VERSION     AF221944.2  GI:14043016
KEYWORDS    .
SOURCE      Ehrlichia chaffeensis.
  ORGANISM  Ehrlichia chaffeensis
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Anaplasmataceae; Ehrlichia.
REFERENCE   1  (bases 1 to 1542)
  AUTHORS   Lee,K.N., Massung,R.F., Padmalayam,I. and Baumstark,B.
  TITLE     Characterization of the ftsZ gene in Ehrlichia chaffeensis, the HGE
            agent and Rickettsia rickettsii
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1542)
  AUTHORS   Lee,K.N., Massung,R.F., Padmalayam,I. and Baumstark,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2000) Centers for Disease Control and Prevention,
            1600 Clifton Rd., Decatur, GA 30329, USA
COMMENT     On May 15, 2001 this sequence version replaced gi:12655827.
FEATURES             Location/Qualifiers
     source          1. .1542
                     /organism="Ehrlichia chaffeensis"
                     /db_xref="taxon:945"
     gene            254. .1519
                     /gene="ftsZ"
     CDS             254. .1519
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein FtsZ"
                     /protein_id="AAK00615.2"
                     /db_xref="GI:14043017"
                     /translation="MSLNICLPDQSLLRPRITVFGVGGAGGNAVNNMIQSNLHGVNFV
                     VANTDAQALELSLSEKKIQLGIGLTKGLGAGSLPEVGRGAAEESINEIIEEISDSNML
                     FITAGMGGGTGTGAAPVIARVAKENKILTIGVVTKPFHFEGAHRMRTAEFGLEELQRY
                     VDTLIVIPNQNLFRIANEKTTFADAFKLADTVLHTGVRGITDLMVMPGLINLDFADIR
                     AIMSEMGKAMMGTGEAEGENRAIAAAEAAISNPLLDNISMKGAKGILINITGGLDMTL
                     FEVDAAANRIREEVDSHANIIFGSTFNKESEGKIRVSVLATGIDNEEVVIQNKSALKD
                     KEAHNDKLSETSSKPFNPLDNEIAYYKPSNPGEDMINSINHTKKQDHNIENQKSNNKI
                     PESAYKMQFNKNEYWDEESFNIPTFLRKK"
BASE COUNT      577 a    232 c    280 g    453 t
ORIGIN

Alignment Scores:
Pred. No.:                3.49e-39       Length:         1542
Score:                    525.00         Matches:        107
Percent Similarity:       75.14%         Conservative:   26
Best Local Similarity:    60.45%         Mismatches:     44
Query Match:              59.32%         Indels:         0
DB:                       1              Gaps:           0

US-09-770-509-2 (1-178) x AF221944 (1-1542)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||   ||||||   |||:::|||||||||||||||   |||||||||
Db        359 TCTAATCTGCACGGCGTTAACTTTGTAGTAGCTAACACAGATGCACAAGCATTAGAACTT 418

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              ||||||:::      ||||||   ||||||      :::|||||||||||||||||||||
Db        419 TCCTTGTCAGAAAAGAAAATTCAGCTAGGTATCGGATTAACAAAAGGACTAGGAGCAGGA 478

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              |||   |||||||:::|||:::   :::||||||:::         :::|||
Db        479 TCACTTCCAGAAGTGGGTAGAGGTGCTGCAGAAGAGTCAATAAATGAGATTATTGAGGAA 538

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              :::   ||||||||||||||||||||||   |||||||||||||||||   ||||||
Db        539 ATCTCAGATAGCAATATGCTGTTCATAACTGCTGGAATGGGTGGAGGAACAGGAACAGGC 598

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              |||||||||||||:::|||   ||||||:::|||      ||||||||:::||||||||
Db        599 GCTGCTCCTGTAATTGCCAGAGTTGCCAAGGAGAACAAAATTTTAACCATAGGAGTTGTT 658

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||   :::|||   |||   |||      |||::::::
Db        659 ACAAAACCTTTCCACTTCGAAGGGGCACATCGAATGAGGACAGCAGAATTTGGGTTAGAA 718

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              ||||||   :::|||||||||||||||||||||:::|||||||||||||||      :::
Db        719 GAATTACAAAGGTATGTTGATACTCTCATAGTAATTCCTAATCAAAATCTATTCAGGATT 778

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              |||::::::   ||||||      :::|||||:::   ||||||   ||||||
Db        779 GCTAATGAAAAAACAACATTTGCAGATGCATTCAAACTTGCAGACACTGTCTTACATACA 838

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||:::||||||||||:::|||   ||||||||||||||||||
Db        839 GGTGTTCGTGGCATCACTGACTTAATGGTAATGCCAGGGCTCATTAATCTA 889

RESULT 5
AF221946
LOCUS       AF221946               1540 bp    DNA     linear   BCT 15-MAY-2001
DEFINITION  Rickettsia rickettsii cell division protein FtsZ (ftsZ) gene,
            complete cds.
ACCESSION   AF221946
```

```
VERSION     AF221946.2  GI:14043019
KEYWORDS    .
SOURCE      Rickettsia rickettsii.
  ORGANISM  Rickettsia rickettsii
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Rickettsieae; Rickettsia; spotted fever group.
REFERENCE   1  (bases 1 to 1540)
  AUTHORS   Lee,K.N., Massung,R.F., Padmalayam,I. and Baumstark,B.
  TITLE     Characterization of the ftsZ gene in Ehrlichia chaffeensis, the HGE
            agent and Rickettsia rickettsii
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1540)
  AUTHORS   Lee,K.N., Massung,R.F., Padmalayam,I. and Baumstark,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2000) Centers for Disease Control and Prevention,
            1600 Clifton Rd., Decatur, GA 30329, USA
COMMENT     On May 15, 2001 this sequence version replaced gi:12655831.
FEATURES             Location/Qualifiers
     source          1. .1540
                     /organism="Rickettsia rickettsii"
                     /db_xref="taxon:783"
     gene            39. .1397
                     /gene="ftsZ"
     CDS             39. .1397
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein FtsZ"
                     /protein_id="AAK00617.1"
                     /db_xref="GI:12655832"
                     /translation="MVLNIKAPENIVLKPTITVFGVGGAGSNAVNNMISANLQGANFV
                     VANTDAQSLEHSLCTNKIQLGVSTTRGLGAGASPEVGALAAQESESEIRNYLENSNMV
                     FITAGMGGGTGTGSAPVIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGLIELQQF
                     VDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINLDFADIK
                     AVMSEMGKAMMGTGEASGEDRAIKAAESAISNPLLDHSSMCGARGVLINITGGSDMTL
                     FEVDNAANRIREEVDNLDANIIFGSTFNPELKGMIRVSVVATGIDADKVPTYKPAIAE
                     TTNIVPEETYNKAIAQPTQIEEMPDFNSYSTKNIEITDSPINQNLIGNEKELGLHANT
                     FNKSEDDSPKPSFLGKIWGSLRASNNQTLERKNVVVSTLDQDNKESDIHDIPAFLRKQ
                     RD"
BASE COUNT      537 a    267 c    305 g    431 t
ORIGIN

Alignment Scores:
Pred. No.:                4.31e-39       Length:          1540
Score:                    524.00         Matches:         100
Percent Similarity:       77.53%         Conservative:    38
Best Local Similarity:    56.18%         Mismatches:      40
Query Match:              59.21%         Indels:          0
DB:                       1              Gaps:            0

US-09-770-509-2 (1-178) x AF221946 (1-1540)

Qy          1 AlaSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
              ::::::   |||:::|||      |||:::||||||||||||||||   |||:::|||
Db        141 AGTGCTAATCTGCAAGGTGCTAATTTTGTAGTAGCTAATACTGATGCACAATCGCTTGAA 200

Qy         21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
                ||||||      :::||||||   ||||||         |||:::|||||||||||
Db        201 CATTCTTTATGCACTAACAAAATACAACTCGGTGTTTCTACGACTAGAGGTCTTGGGGCA 260

Qy         41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
              |||:::   ||||||:::|||      :::|||::::::   :::   :::|||:::
Db        261 GGAGCTTCTCCTGAGGTTGGAGCACTTGCTGCTCAAGAATCAGAAAGCGAAATTCGTAAT 320

Qy         61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
                |||::::::|||||||||:::|||||||||   ||||||||||||||||||   |||
Db        321 TACCTAGAAAATAGCAATATGGTATTTATTACGGCAGGTATGGGCGGTGGTACAGGTACC 380

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
              |||:::|||||||||:::|||   :::|||:::||||||||||||||||||||||||||
Db        381 GGTTCTGCACCGGTTATTGCACGCATTGCTAAAGAACTAGGCATCCTTACGGTTGGGGTA 440

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
              |||:::   ||||||      ||||||   ::::||   :::   |||:::   |||:::
Db        441 GTAACTAAACCTTTCCATTTTGAAGGCGGTCATCGTATGAAAACTGCCGATAAAGGACTT 500

Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
                ||||||   :::::|||||||||||||||||:::|||||||||||||||
Db        501 ATTGAATTACAGCAATTCGTTGATACTTTAATTGTAATACCGAACCAAAATCTATTCCGT 560

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
              :::|||::::::   ||||||      :::|||||||:::   ||||||||||||||||
Db        561 ATTGCTAATGAACAAACGACGTTTGCCGATGCTTTCAAAATGGCAGATGACGTATTACAT 620

Qy        161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                ||||||:::|||||||||||||||||:::::   ||||||||||||||||||
Db        621 GCAGGCGTTAGAGGAGTAACGGATTTAATGATTATGCCGGGACTTATTAATCTA 674

RESULT 6
AE008653
LOCUS       AE008653                8328 bp    DNA     linear   BCT 14-SEP-2001
DEFINITION  Rickettsia conorii Malish 7, section 85 of 114 of the complete
            genome.
ACCESSION   AE008653 AE006914
VERSION     AE008653.1  GI:15620130
KEYWORDS    .
SOURCE      Rickettsia conorii.
  ORGANISM  Rickettsia conorii
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Rickettsieae; Rickettsia; spotted fever group.
REFERENCE   1  (sites)
  AUTHORS   Ogata,H., Audic,S., Barbe,V., Artiguenave,F., Fournier,P.E.,
            Raoult,D. and Claverie,J.M.
  TITLE     Selfish DNA in protein-coding genes of Rickettsia
  JOURNAL   Science 290 (5490), 347-350 (2000)
  MEDLINE   20485642
   PUBMED   11030655
REFERENCE   2  (sites)
  AUTHORS   Ogata,H., Audic,S. and Claverie,J.-M.
  TITLE     Selfish DNA and the origin of genes
  JOURNAL   Science 291 (5502), 252-253 (2001)
REFERENCE   3  (bases 1 to 8328)
  AUTHORS   Ogata,H., Audic,S., Renesto-Audiffren,P., Fournier,P.-E., Barbe,V.,
            Samson,D., Roux,V., Cossart,P., Weissenbach,J., Claverie,J.-M. and
            Raoult,D.
  TITLE     Mechanisms of evolution in Rickettsia conorii and R. prowazekii
  JOURNAL   Science 293 (5537), 2093-2098 (2001)
  MEDLINE   21442074
   PUBMED   11557893
REFERENCE   4  (bases 1 to 8328)
  AUTHORS   Ogata,H., Audic,S., Renesto-Audiffren,P., Fournier,P.-E., Barbe,V.,
            Samson,D., Roux,V., Cossart,P., Weissenbach,J., Claverie,J.-M. and
            Raoult,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-2001) Unite des Rickettsies, Faculte de Medecine,
            CNRS UMR 6020, Universite de la Mediterranee, 27 boulevard Jean
            Moulin, Marseille Cedex 05 13385, France
COMMENT     A public version of R. conorii genome database is accessible at
            http://igs-server.cnrs-mrs.fr/. The database intends to provide
            updated data. Annotation of the genome is an ongoing task whose
            goal is to make the genome sequence more useful. Comments to the
            authors are appreciated.
FEATURES             Location/Qualifiers
     source          1. .8328
                     /organism="Rickettsia conorii"
                     /strain="Malish 7"
                     /db_xref="taxon:781"
     gene            203. .1561
                     /gene="ftsZ"
                     /note="RC1015"
     CDS             203. .1561
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein ftsZ"
                     /protein_id="AAL03553.1"
```

```
                /db_xref="GI:15620131"
                /translation="MVLNIKAPENIVLKPTITVFGVGGAGSNAVNNMISANLQGANFV
                VANTDAQSLEHSLCTNKIQLGVSTTRGLGAGASPEVGALAAQESESEIRNYLENSNMV
                FITAGMGGGTGTGSAPVIARIAKELGILTVGVVTKPFHFEGGHRMKTADKGLIELQQF
                VDTLIVIPNQNLFRIANEQTTFADAFKMADDVLHAGVRGVTDLMIMPGLINLDFADIK
                AVMSEMGKAMMGTGEASGEDRAIKAAESAISNPLLDHSSMCGARGVLINITGGSDMTL
                FEVDNAANRIREEVDNLDANIIFGSTFNPELKGMIRVSVVATGIDADKVPTYKPAIAE
                TTNIVPEETYNKAIAQPTQIEEMPDFNSYSTENIEITDSPINQNFIGNEKELGLHANT
                FNKSEDDSPKPSFLGKIWGSLRASNNQTLERKNVVVSTLDQDNKESDIHDIPAFLRKK
                RD"
gene            1835. .2473
                /gene="RC1016"
CDS             1835. .2473
                /gene="RC1016"
                /codon_start=1
                /transl_table=11
                /product="unknown"
                /protein_id="AAL03554.1"
                /db_xref="GI:15620132"
                /translation="MEDKPVYSLDIDGATYYAALDFTELNQNNMLNTRFSATSLDFHL
                IETLPNFSSDKQKYKAELTFTELKQNNFAKIYLEKQQYNSQSILISNIQNNHNNYILF
                PKIINLEEQVTIDILNNLLLLGTYNKCNFFESNNGSEIYFRLPNNERLAMKVFSLNNQ
                PILKENKFENSILFLSTKINENGDVDYSEGFSNSQWYNKCFLYNSSRRDNGF"
gene            2660. .2866
                /gene="RC1017"
CDS             2660. .2866
                /gene="RC1017"
                /codon_start=1
                /transl_table=11
                /product="similarity to ecotin precursor"
                /protein_id="AAL03555.1"
                /db_xref="GI:15620133"
                /translation="MQATKNAMQDCNRVWFGGKLETKTLEGWGYNYYIIDQVSDHPAS
                TMMACPNVKATIQTVSVFLGDETF"
repeat_region   3006. .3031
                /note="REP01, repeated element"
gene            complement(3084. .3656)
                /gene="RC1018"
CDS             complement(3084. .3656)
                /gene="RC1018"
                /codon_start=1
                /transl_table=11
                /product="unknown"
                /protein_id="AAL03556.1"
                /db_xref="GI:15620134"
                /translation="MFIQTEDTPNPDAITFFPGQEISSDQPVFFSELAEVKGRSKLAE
                SLFYINNVKSVFFGSDFITVTKQAEGNWQVIKPEILMVIMDHFVSGFPVFEENTKADN
                VNHNLDGLSEIEKQIVEIIETRVRPSVAQDGGDIIYKGFENGVVKLALRGACLGCPSS
                TITLKNGIESMLKHFVPEVQEVEAVEEGFE"
repeat_region   complement(3852. .3880)
                /note="REP01, repeated element"
repeat_region   3920. .3957
                /note="REP08, short repeated element"
gene            complement(4004. .5068)
                /gene="ampG2"
                /note="RC1019"
CDS             complement(4004. .5068)
                /gene="ampG2"
                /codon_start=1
                /transl_table=11
                /product="ampG protein"
                /protein_id="AAL03557.1"
                /db_xref="GI:15620135"
                /translation="MNFIHIFKFLWGPLLEKVSFAPLSKRGYKYCLIIALVSCIFCVY
                VLTNFNPNTHFIPFALCLVAVAFFSSIYDMLLQSSQMLLITNKNWGISEAACTTGFRI
                GILIAGSGALYLSTIISWQDVYRAMAILCIPSLLLIIIYPLKFKDKIIINDFDRFWHA
                FYDFVKKPKWLIIVSFMLLYRLQDNFLSVMPNMFYLDIGYTKKDLALGYKAFGMCATI
                LGGFIGGFLCRKYEYSYLLKRVLIYHALSSLSFLFLYFYNRDITSLYIAVFFQEFTKG
                LTMSPFFSYQLRCCSAKYCITQIALITSIAYISTILFGSISGYAATYLGWTYFFIVAG
                FCFIPAYILIRYLPPYVISS"
repeat_region   5303. .5402
                /note="REP01, repeated element"
gene            complement(5424. .6668)
                /gene="rhlE"
                /note="RC1020"
CDS             complement(5424. .6668)
                /gene="rhlE"
                /codon_start=1
                /transl_table=11
                /product="ATP-dependent RNA helicase RhlE"
                /protein_id="AAL03558.1"
                /db_xref="GI:15620136"
                /translation="MDYIMKNFNLSEELIIALETMNITEPTEIQKQSIPVAMAGSDIL
                ASSQTGSGKTLAYLLPLIDSFIKNKTTALILVPTRELATQIHSTLNKVTTSYKINSAV
                LIGGEPMPKQFIQLKKNPKVIIGTPGRIIDHLNRGSLKIDRIGITVLDEMDRMLDMGM
                KEQLEEINKFLPEKRQVLMFSATMPKHIIAVSQKYLNNPVRITVGATNKAAAEIKQES
                MHVSDKEKFSALTKQLGNREGSVIIFVKTKRSADQLAKMLKYENHTAEAIHGDLSQRQ
                RERVILSFRKSNHRIMVATDVAARGLDIPHTQHVINYDLPMCPEDYLHRIGRTGRAGA
                TGHALSFISPDDVIRWRAIDRLVNKGESTPRSEFRSDKNNRKRSFGKRAGGDGKKFNA
                FDSSSRKRTFDGNSYGNKRKKVS"
repeat_region   complement(6808. .6836)
                /note="REP05, Rickettsia palindromic element 1"
gene            complement(7165. .7377)
                /gene="cspA"
                /note="RC1021"
CDS             complement(7165. .7377)
                /gene="cspA"
                /codon_start=1
                /transl_table=11
                /product="cold shock-like protein"
                /protein_id="AAL03559.1"
                /db_xref="GI:15620137"
                /translation="MATNIVGKVKWYNSTKNFGFIEQDNGGKDVFVHKSAVDAAGLHS
                LEEGQDVIFDLEEKQGKAYAVNLRIK"
gene            7756. .7838
                /gene="RCRNA30"
tRNA            7756. .7838
                /gene="RCRNA30"
                /product="tRNA-Leu"
BASE COUNT     2993 a   1445 c   1263 g   2627 t
ORIGIN

Alignment Scores:
Pred. No.:                2.99e-38        Length:         8328
Score:                    524.00          Matches:        100
Percent Similarity:       77.53%          Conservative:   38
Best Local Similarity:    56.18%          Mismatches:     40
Query Match:              59.21%          Indels:         0
DB:                       1               Gaps:           0

US-09-770-509-2 (1-178) x AE008653 (1-8328)

Qy          1 AlaSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
              ::::::   |||:::|||      |||:::|||||||||||||||   |||:::|||
Db        305 AGTGCTAATCTGCAAGGTGCTAATTTTGTAGTAGCTAATACTGATGCACAATCGCTTGAA 364

Qy         21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
                ||||||      :::||||||   ||||||         |||:::||||||||||||
Db        365 CATTCTTTATGCACTAACAAAATACAACTCGGTGTTTCTACGACTAGAGGTCTTGGGGCA 424

Qy         41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
              |||:::   ||||||:::|||      :::|||:::::   :::   :::|||:::
Db        425 GGAGCTTCTCCTGAGGTTGGAGCACTTGCTGCTCAAGAATCAGAAAGCGAAATTCGTAAT 484

Qy         61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
              |||:::::|||||||||:::|||||||||   |||||||||||||||||  |||
Db        485 TACCTAGAAAATAGCAATATGGTATTTATTACGGCAGGTATGGGCGGTGGTACAGGTACC 544

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
              |||:::|||||||||:::|||   :::|||:::|||||||||||||||||||||||||||
Db        545 GGTTCTGCACCGGTTATTGCACGCATTGCTAAAGAACTAGGCATCCTTACGGTTGGGGTA 604

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
              |||:::   ||||||      ||||||   :::|||   :::   |||:::   |||:::
Db        605 GTAACTAAACCTTTCCATTTTGAAGGCGGTCATCGTATGAAAACTGCCGATAAAGGACTT 664
```

```
Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
              ||||||   :::::|||||||||||||||||||:::|||||||||||||||
Db        665 ATTGAATTACAGCAATTCGTTGATACTTTAATTGTAATACCGAACCAAAATCTATTCCGT 724

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
              :::|||::::::   ||||||      :::||||||:::   ||||||||||||||||
Db        725 ATTGCTAATGAACAAACGACGTTTGCCGATGCTTTCAAAATGGCAGATGACGTATTACAT 784

Qy        161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||||||||||||||||:::::   ||||||||||||||||||
Db        785 GCAGGCGTTAGAGGAGTAACGGATTTAATGATTATGCCGGGACTTATTAATCTT 838

RESULT 7
AF221945
LOCUS       AF221945               1474 bp    DNA     linear   BCT 15-MAY-2001
DEFINITION  Ehrlichia sp. 'HGE agent' cell division protein FtsZ (ftsZ) gene,
            complete cds.
ACCESSION   AF221945
VERSION     AF221945.2  GI:14043018
KEYWORDS    .
SOURCE      Anaplasma phagocytophilum.
  ORGANISM  Anaplasma phagocytophilum
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Anaplasmataceae; Anaplasma; phagocytophilum group.
REFERENCE   1  (bases 1 to 1474)
  AUTHORS   Lee,K.N., Massung,R.F., Padmalayam,I. and Baumstark,B.
  TITLE     Characterization of the ftsZ gene in Ehrlichia chaffeensis, the HGE
            agent and Rickettsia rickettsii
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1474)
  AUTHORS   Lee,K.N., Massung,R.F., Padmalayam,I. and Baumstark,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2000) Centers for Disease Control and Prevention,
            1600 Clifton Rd., Decatur, GA 30329, USA
COMMENT     On May 15, 2001 this sequence version replaced gi:12655829.
FEATURES             Location/Qualifiers
     source          1. .1474
                     /organism="Anaplasma phagocytophilum"
                     /db_xref="taxon:948"
                     /note="human granulocytic ehrlichiosis agent"
     gene            178. .1380
                     /gene="ftsZ"
     CDS             178. .1380
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein FtsZ"
                     /protein_id="AAK00616.1"
                     /db_xref="GI:12655830"
                     /translation="MSLHVYLPDQSVIRPRITVLGVGGAGGNAVNNMIQSCLQGVNFV
                     VANTDAQALDCSLSEKKIQLGMNLTKGLGAGSLPEIGRGAAEESIEEIIAEISDSNML
                     FITAGMGGGTGTGAAPVIAKAAKDSKILTVGVVTRPFHFEGAHRMKTAEYGLEELQKH
                     VDTLIVIPNQNLFRIANENTTFADAFKLADTVLHTGVRGITDLMVMPGLINLDFADVK
                     VVMSEMGKAMMGTGEAEGEHRAVAAAEAAISNPLLDNISMKGARGILINITGGMDMTL
                     FEVDAAANRIREEVDEEANIIFGSTFDENSAGRIRVSVLATGIDSTHTSNPKGRMTRG
                     SESDLGVSDVGGDSARVGVSGYARAGQESSGAPAEADLFSRGSGNRNSSLNWNSKSFD
                     VPAFLRKK"
BASE COUNT      382 a    255 c    416 g    421 t
ORIGIN

Alignment Scores:
Pred. No.:              7.73e-39       Length:         1474
Score:                  521.00         Matches:        105
Percent Similarity:     76.27%         Conservative:   30
Best Local Similarity:  59.32%         Mismatches:     42
Query Match:            58.87%         Indels:         0
DB:                     1              Gaps:           0

US-09-770-509-2 (1-178) x AF221945 (1-1474)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::|||||||   |||:::|||||||||||||||||   |||||||||
Db        283 TCTTGTTTGCAGGGCGTTAATTTTGTTGTGGCAAATACTGACGCTCAGGCGTTGGATTGT 342

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||||||:::       ||||||   ||||||   ::::::|||||||||||||||||||||
Db        343 TCATTATCAGAGAAGAAGATCCAGTTGGGAATGAACCTCACAAAAGGTTTGGGGGCTGGT 402

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              |||   |||||||:::|||:::   :::|||||||:::          :::|||
Db        403 TCTCTGCCTGAGATTGGTCGTGGAGCGGCTGAGGAATCAATAGAGGAGATAATCGCAGAG 462

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              :::   ||||||||||||||||||||||||   ||||||||||||||||||   ||||||
Db        463 ATCTCAGATAGCAATATGCTATTTATAACTGCGGGTATGGGTGGAGGCACTGGAACTGGT 522

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||:::|||      |||::::::      ||||||||||||||||||||
Db        523 GCAGCTCCTGTTATCGCTAAGGCTGCAAAAGACAGCAAAATTCTAACTGTGGGTGTGGTC 582

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||   :::|||   :::   |||      |||::::::
Db        583 ACTAGGCCATTTCATTTCGAAGGTGCCCACAGGATGAAAACTGCGGAATACGGTCTGGAG 642

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              ||||||   |||:::|||||||||||||||||:::|||||||||||||||      :::
Db        643 GAGCTGCAAAAGCATGTTGATACCTTGATCGTAATTCCTAATCAGAATTTGTTTCGTATA 702

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              |||:::::::::|||||      :::|||||||:::   ||||||   ||||||
Db        703 GCTAACGAGAACACTACTTTCGCTGACGCGTTCAAGTTGGCCGACACTGTGTTGCATACA 762

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              |||||||:::|||:::|||||||||:::|||   ||||||||||||||||
Db        763 GGCGTGCGTGGTATCACAGACTTAATGGTTATGCCAGGGTTGATAAACCTA 813

RESULT 8
AF067822
LOCUS       AF067822               1202 bp    DNA     linear   BCT 18-AUG-1998
DEFINITION  Clostridium lentocellum cell division protein (ftsZ) gene, complete
            cds.
ACCESSION   AF067822
VERSION     AF067822.1  GI:3426307
KEYWORDS    .
SOURCE      Clostridium lentocellum.
  ORGANISM  Clostridium lentocellum
            Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
            Clostridium.
REFERENCE   1  (bases 1 to 1202)
  AUTHORS   Angert,E.R. and Losick,R.M.
  TITLE     Propagation by sporulation in the guinea pig symbiont Metabacterium
            polyspora
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (17), 10218-10223 (1998)
  MEDLINE   98374332
   PUBMED   9707627
REFERENCE   2  (bases 1 to 1202)
  AUTHORS   Angert,E.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-1998) MCB, Harvard University, 16 Divinity Ave.,
            Cambridge, MA 02138, USA
FEATURES             Location/Qualifiers
     source          1. .1202
                     /organism="Clostridium lentocellum"
                     /strain="ATCC49066"
                     /db_xref="ATCC:49066"
                     /db_xref="taxon:29360"
     gene            45. .1157
                     /gene="ftsZ"
     CDS             45. .1157
                     /gene="ftsZ"
                     /function="involved in prokaryotic cell division"
                     /note="FtsZ"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein"
                     /protein_id="AAC32265.1"
```

```
                     /db_xref="GI:3426308"
                     /translation="MLEIWENELQGAQIKVIGVGGGGNNAVDRMIEKGLEGVEFITVN
                     TDHQALARSGAPAKIQIGEKMTRGLGAGANPEIGTKSAEESREEILTAIKGADMLFIT
                     AGMGGGTGTGAAPVIASIAKEEGILTVGVVTKPFSFEGRKRMINAEKGIAELKQNVDT
                     LVVIPNDKILQVIDKKTTMVDAFSKADDVLQQGVQGITDLISNPGIINLDFADVRTIM
                     NNKGVAHMGIGRATGENRAEEAVKYAISSPLLDTSIDGARCVLVNMCGGESLGLMEAN
                     VGMGLIREAVDPDAEIIFGTSINENLGEEIIITVIATDFQNHDVSLNTFKPVQATKEE
                     TPVQTQKATTAHVEETKFTPIKDIQVEIPQFLRRNR"
BASE COUNT      428 a    167 c    272 g    335 t
ORIGIN

Alignment Scores:
Pred. No.:                7.55e-39        Length:          1202
Score:                    520.00          Matches:         103
Percent Similarity:       76.00%          Conservative:    30
Best Local Similarity:    58.86%          Mismatches:      42
Query Match:              58.76%          Indels:          0
DB:                       1               Gaps:            0

US-09-770-509-2 (1-178) x AF067822 (1-1202)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ||||||||||||||||||||      |||||||||  |||||||||   ||||||
Db        147 CTTGAAGGTGTTGAATTTATTACTGTTAATACAGATCATCAAGCACTGGCTAGATCAGGT 206

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              ||||||   ||||||   :::|||:::   :::|||:::|||||||||||||||:::
Db        207 GCACCAGCCAAAATTCAGATTGGTGAAAAAATGACACGTGGACTTGGAGCAGGCGCTAAT 266

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||   :::||||||||:::   :::   :::|||         ::::::
Db        267 CCAGAAATTGGTACAAAATCAGCAGAGGAAAGTCGTGAGGAGATACTAACAGCTATTAAA 326

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||||||||||||||   ||||||||||||||||||   ||||||||||||
Db        327 GGCGCAGATATGCTATTTATCACAGCCGGAATGGGTGGTGGTACAGGTACAGGTGCAGCA 386

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::||||||:::|||:::|||   |||||||||||||||||||||||:::
Db        387 CCAGTTATTGCGAGTATTGCAAAAGAAGAAGGTATTTTAACAGTTGGTGTCGTGACTAAG 446

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||  ||||         |||      ||:::   ||||||
Db        447 CCATTTAGTTTTGAAGGTCGCAAAAGAATGATTAATGCAGAGAAAGGTATTGCAGAGCTT 506

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::|||:::||||||      :::|||   :::   |||
Db        507 AAGCAAAATGTAGATACCTTAGTTGTTATCCCAAATGATAAAATTTTACAAGTTATTGAT 566

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||||||||:::::|||||      ||||||||||||||   :::||||||
Db        567 AAGAAGACAACGATGGTGGACGCATTTAGTAAGGCCGACGATGTATTACAACAAGGTGTA 626

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::||||||||||||      |||||||:::||||||||
Db        627 CAGGGGATTACAGATTTAATTTCAAACCCAGGTATTATTAACCTT 671

RESULT 9
AE013119/c
LOCUS       AE013119               11683 bp    DNA     linear   BCT 09-MAY-2002
DEFINITION  Thermoanaerobacter tengcongensis strain MB4T, section 146 of 244 of
            the complete genome.
ACCESSION   AE013119 AE008691
VERSION     AE013119.1  GI:20516647
KEYWORDS    .
SOURCE      Thermoanaerobacter tengcongensis.
  ORGANISM  Thermoanaerobacter tengcongensis
            Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
            Thermoanaerobacteriaceae; Thermoanaerobacter.
REFERENCE   1  (bases 1 to 11683)
  AUTHORS   Bao,Q., Tian,Y., Li,W., Xu,Z., Xuan,Z., Hu,S., Dong,W., Yang,J.,
            Chen,Y., Xue,Y., Xu,Y., Lai,X., Huang,L., Dong,X., Ma,Y., Ling,L.,
            Tan,H., Chen,R., Wang,J., Yu,J. and Yang,H.
  TITLE     A Complete Sequence of the T. tengcongensis Genome
  JOURNAL   Genome Res. 12 (5), 689-700 (2002)
  MEDLINE   21992816
   PUBMED   11997336
REFERENCE   2  (bases 1 to 11683)
  AUTHORS   Bao,Q., Xu,Z., Hu,S., Dong,W., Chen,Y., Wang,J., Yu,J. and Yang,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-OCT-2001) Beijing Genomics Institute/Genomics and
            Bioinformatics Center, Institute of Genetics and Development,
            Chinese Academy of Sciences, Beijing Airport Industrial Zone B6,
            Beijing 101300, China
REFERENCE   3  (bases 1 to 11683)
  AUTHORS   Li,W., Xuan,Z., Yang,J., Ling,L. and Chen,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-OCT-2001) Institute of Biophysics, Chinese Academy of
            Sciences, Beijing 100101, China
REFERENCE   4  (bases 1 to 11683)
  AUTHORS   Tian,Y., Xue,Y., Xu,Y., Lai,X., Huang,L., Dong,X., Ma,Y. and Tan,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-OCT-2001) Institute of Microbiology, Chinese Academy
            of Sciences, Beijing 100080, China
FEATURES             Location/Qualifiers
     source          1. .11683
                     /organism="Thermoanaerobacter tengcongensis"
                     /strain="MB4T"
                     /db_xref="taxon:119072"
     terminator      complement(88. .107)
                     /note="putative rho-independent transcription terminator"
     gene            complement(109. .1195)
                     /gene="FtsZ"
                     /note="TTE1639"
     CDS             complement(109. .1182)
                     /gene="FtsZ"
                     /note="Best Blastp hit =  gi|10175178|dbj|BAB06277.1|
                     (AP001515) cell-division initiation protein (septum
                     formation) [Bacillus halodurans], score 352, E-value
                     4.00E-96"
                     /codon_start=1
                     /transl_table=11
                     /product="Cell division GTPase"
                     /protein_id="AAM24841.1"
                     /db_xref="GI:20516648"
                     /translation="MIGLETDVQQFAAIKVIGVGGGGGNAVNRMVEAGVKGVEFIAIN
                     TDKQALSLSKAETKIQIGEKLTKGLGAGANPEIGKKAAEESREEIERAIKGADMIFIT
                     AGMGGGTGTGAAPVVAEIAKELGILTVGVVTKPFTFEGRKRMAQAEMGIEDLKKYVDA
                     LITIPNDRLLQVVEKKTSMLDAFKLADDVLRQGVQGISDLIAVPGLVNVDFADVKTIM
                     VNTGLAHMGIGIASGENKATEAAKQAIHSPLLETSIEGSKGILLNIAGGPNLTIFEVN
                     EAANFIYEAADPDANIIFGAVIDEALEDQIRITVIATGFEKNSKAKEEPKKKEVKGVD
                     VKLEDILDNDDLDIPTFLRRGRR"
     misc_feature    complement(526. .1011)
                     /gene="FtsZ"
                     /note="Pfam match to entry tubulin, Tubulin/FtsZ family,
                     score 95.9, E-value 8.10E-25"
     RBS             complement(1191. .1195)
                     /gene="FtsZ"
     gene            complement(1197. .2437)
                     /gene="FtsA2"
                     /note="TTE1640"
     CDS             complement(1197. .2423)
                     /gene="FtsA2"
                     /note="Best Blastp hit =  gi|120567|sp|P28264|FTSA_BACSU
                     CELL DIVISION PROTEIN FTSA gi|2126939|pir||I39847
                     cell-division protein (septum formation) ftsA - Bacillus
                     subtilis gi|142940|gb|AAA22456.1| (M22630) ftsA [Bacillus
                     subtilis] gi|2633899|emb|CAB13401.1| (Z99111)
                     cell-division protein [Bacillus subtilis], score 306,
                     E-value 2.00E-82"
                     /codon_start=1
                     /transl_table=11
                     /product="predicted ATPases of the HSP70 class involved in
                     cell division"
                     /protein_id="AAM24842.1"
                     /db_xref="GI:20516649"
```

/translation="MGDLIVGLDIGTSKVCTIIGEGDKTGELHIVGIGYYPSTGVKKG
IIVDIDETAYSIQKSVEQAERMANRKISSVYLKIYGGLTTIYKNNGVVAVSREDREIT
RQDVERVLQAAKIIALPSDKQIIDVIPLEYIVDGYGEIKDPVGMSGIRLEVNAAVVTG
SVTAIQNMEKCVRKAGLEIDGIIVGPLATSEAVLLKDEKELGVALIDVGAGVTDISVF
KNGGLIYSSMIAVGGWHITNDLSVGLKISFEEAENIKKKYGTLEKVDPDRLEPIKIAS
LAGKSKTTTDINEIADIIEARVSELLTLVYERLEEAGVLEDIVTNVVITGGGISFLKG
SVELAQKIFDRNVRIGSPQNIGVATPIYSAGVGVVKYVYNNRRFMYPSTSTDKKEKKN
SRLLDKLKEIFSDFWA"

misc_feature complement(1203. .2342)
/gene="FtsA2"
/note="Pfam match to entry FtsA, Cell division protein
FtsA, score 452.4, E-value 3.80E-132"

RBS complement(2433. .2437)
/gene="FtsA2"

gene complement(2448. .2807)
/gene="TTE1641"

CDS complement(2448. .2792)
/gene="TTE1641"
/note="Best Blastp hit = gi|10175180|dbj|BAB06279.1|
(AP001515) small basic protein [Bacillus halodurans],
score 100, E-value 5.00E-21"
/codon_start=1
/transl_table=11
/product="small basic protein"
/protein_id="AAM24843.1"
/db_xref="GI:20516650"
/translation="MAYIIALSLLLGLIIGFVLPINIPFSYAPYLSVAILAALDSVFG
GIRASFEEKFDNLIFITGFFGNIILAGGLAYIGDVLGVPIYLAAVFVFGSRLFSNFAY
IRRYLIDRMRRR"

gene complement(2792. .3566)
/gene="TTE1642"

CDS complement(2792. .3520)
/gene="TTE1642"
/note="Best Blastp hit = gi|1075848|pir||C43727 conserved
hypothetical protein ylxW - Bacillus subtilis
gi|1129091|gb|AAA83971.1| (M31827) ORF4; putative
[Bacillus subtilis] gi|2633896|emb|CAB13398.1| (Z99111)
similar to hypothetical proteins [Bacillus subtilis],
score 126, E-value 2.00E-28"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM24844.1"
/db_xref="GI:20516651"
/translation="MKKGFSIYIISIALVSLVLGIMLSTQFKIVKKGGVVTAQRAEELA
TQLKQLQSEKEALLKQISELEAKINAYEDSASKNSVITQTLKSEVEKYKELAGLTDLQ
GPGVIVTVNDSKQPVPPGQDPNLFLVHDEDLLRIVNELRAAGAEAISLNDQRLIATSE
IRCVGPTVVVNSVRFAAPYVIKAIGDPDTLEAALKLKGGVVDSLATWGIEVNIKKSDK
IVIKKYDGVLKFRYAKPLVEGEKK"

RBS complement(2803. .2807)
/gene="TTE1641"

gene complement(3536. .4245)
/gene="FtsQ"
/note="TTE1643"

CDS complement(3536. .4234)
/gene="FtsQ"
/codon_start=1
/transl_table=11
/product="Cell division septal protein"
/protein_id="AAM24845.1"
/db_xref="GI:20516652"
/translation="MNKRAFRVIFFLLILAVLSYVFAFHTNYFKIKSIKVVGNQILSY
NDIKEISKIQAGTNIFKVNPKQVEKNLLENPYIKECKVKILYPNRVEIFVEERRVVAQ
VRYKSDYLKIDKEGVIVEKGSFTPGLLLIEGIKVERYQIGKKLNGNFDKTLLSEVLGL
IESKSNYYALRYIKEDEVELLSKQGINIFLKNPADVNYSFKFAELILKDLLKKGYKSG
KIEIIGDGNAVFMQ"

RBS complement(3562. .3566)
/gene="TTE1642"

RBS complement(4241. .4245)
/gene="FtsQ"

gene complement(4291. .5556)
/gene="MurA2"
/note="TTE1644"

CDS complement(4291. .5544)
/gene="MurA2"
/note="Best Blastp hit = gi|7674128|sp|P70965|MUA1_BACSU
UDP-N-ACETYLGLUCOSAMINE 1-CARBOXYVINYLTRANSFERASE 1
(ENOYLPYRUVATE TRANSFERASE) (UDP-N-ACETYLGLUCOSAMINE
ENOLPYRUVYL TRANSFERASE) (EPT) gi|7448125|pir||A69662
UDP-N-acetylglucosamine 1-carboxyvinyltransferase murA -
Bacillus subtilis gi|1648861|emb|CAB03688.1| (Z81356)
UDP-N-acetylglucosamine 1-carboxyvinyltransferase
[Bacillus subtilis] gi|2636201|emb|CAB15693.1| (Z99122)
UDP-N-acetylglucosamine 1-carboxyvinyltransferase
[Bacillus subtilis], score 367, E-value 1.00E-100"
/codon_start=1
/transl_table=11
/product="UDP-N-acetylglucosamine enolpyruvyl transferase"
/protein_id="AAM24846.1"
/db_xref="GI:20516653"
/translation="MGKFKINGGKRLYGEVEVHGAKNSILPILAATILNEGVSVIHNC
PRLKDVDSMIEILEHIGCKVSFSGRDIVVDARDVKDSEIPDNLMRTMRSSIFLMGALI
ARNKKAFISFPGGCDIGHRPIDLHLKGLKKLGVEIEESYGYIRCKGVRVRGNEIHLDL
PSVGATENIMLAATLADGITVIRNAAKEPEIEDLQNFLNSMGARITGAGTNTIVIEGV
KKLHDTEYTIIPDRIVAGTYLCAAAMTRGELTVVKALKEHLEPLISKLKETGCELKTG
NDYIKITCNKRPKAVDMIVTLPYPGFPTDLQPQIVSVLSIAEGTSIVTETIFDNRFKY
TEELVRMGADIKVEGRVAVIRGVDKITGAKVLAKDLRGGVALVIAGLGAEGTTVVEGA
EHIDRGYESLEKALKSVGADIVRIM"

misc_feature complement(4327. .5529)
/gene="MurA2"
/note="Pfam match to entry EPSP_syntase, EPSP synthase
(3-phosphoshikimate 1-carboxyvinyltransferase), score
407.8, E-value 1.00E-118"

RBS complement(5552. .5556)
/gene="MurA2"

gene complement(5618. .6727)
/gene="MurG"
/note="TTE1645"

CDS complement(5618. .6712)
/gene="MurG"
/note="Best Blastp hit = gi|10175185|dbj|BAB06284.1|
(AP001515) UDP-N-acetylglucosamine-N-acetylmuramyl-
(pentapeptide)pyrophos [Bacillus halodurans], score 306,

Alignment Scores:

Pred. No.:	1.03e-37	Length:	11683
Score:	520.00	Matches:	99
Percent Similarity:	77.14%	Conservative:	36
Best Local Similarity:	56.57%	Mismatches:	40
Query Match:	58.76%	Indels:	0
DB:	1	Gaps:	0

US-09-770-509-2 (1-178) x AE013119 (1-11683)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||      |||||||||   |||||||||      |||
Db       1080 GTTAAAGGTGTGGAATTTATAGCCATCAATACAGACAAGCAAGCTCTTTCTTTGTCAAAA 1021

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      ||||||   :::|||:::   :::||||||||||||||||||||||:::
Db       1020 GCTGAAACTAAAATCCAGATAGGAGAAAAATTGACAAAAGGTCTGGGGGCTGGTGCCAAT 961

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||||:::::|||||:::   :::   :::|||:::|||   ::::::
Db        960 CCTGAAATTGGTAAAAAAGCGGCAGAAGAGTCGAGGGAAGAGATCGAAAGAGCGATAAAA 901

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::||||||   ||||||||||||||||   ||||||||||
Db        900 GGTGCTGACATGATATTCATAACAGCGGGCATGGGAGGCGGCACAGGTACAGGAGCAGCT 841

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||||||||||||||||||||||||:::
Db        840 CCAGTTGTAGCTGAAATTGCAAAGGAACTTGGTATTTTAACTGTAGGCGTTGTGACTAAA 781

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||      |||:::::::::|||
```

```
Db        780 CCCTTCACTTTTGAAGGAAGGAAGAGAATGGCGCAGGCAGAGATGGGAATAGAAGATTTA 721

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143

Db        720 AAGAAGTACGTAGATGCACTCATAACGATCCCTAATGATAGATTATTACAGGTAGTGGAG 661

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163

Db        660 AAGAAGACATCTATGCTAGATGCTTTTAAGCTGGCAGATGACGTTTTAAGGCAGGGCGTT 601

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178

Db        600 CAGGGAATTTCTGATTTAATAGCTGTCCCGGGGCTTGTGAATGTG 556

RESULT 10
RPXX03
LOCUS       RPXX03                279110 bp    DNA     linear   BCT 11-NOV-1998
DEFINITION  Rickettsia prowazekii strain Madrid E, complete genome; segment
            3/4.
ACCESSION   AJ235272 AJ235269
VERSION     AJ235272.1  GI:3861033
KEYWORDS    complete genome.
SOURCE      Rickettsia prowazekii.
  ORGANISM  Rickettsia prowazekii
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Rickettsieae; Rickettsia; typhus group.
REFERENCE   1  (bases 1 to 279110)
  AUTHORS   Andersson,S.G., Zomorodipour,A., Andersson,J.O.,
            Sicheritz-Ponten,T., Alsmark,U.C., Podowski,R.M., Naslund,A.K.,
            Eriksson,A.S., Winkler,H.H. and Kurland,C.G.
  TITLE     The genome sequence of Rickettsia prowazekii and the origin of
            mitochondria
  JOURNAL   Nature 396 (6707), 133-140 (1998)
  MEDLINE   99039499
   PUBMED   9823893
REFERENCE   2  (bases 1 to 279110)
  AUTHORS   Andersson,S.G.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-NOV-1998) S.G.E. Andersson,
            Siv.Andersson@molbio.uu.se, Dept. of Molecular Biology, University
            of Uppsala, Husargatan 3, Uppsala, S-751 24, SWEDEN
FEATURES             Location/Qualifiers
     source          1. .279110
                     /organism="Rickettsia prowazekii"
                     /strain="Madrid E"
                     /db_xref="taxon:782"
     gene            63. .1046
                     /gene="RP479"
     CDS             63. .1046
                     /gene="RP479"
                     /codon_start=1
                     /transl_table=11
                     /product="OCTAPRENYL-DIPHOSPHATE SYNTHASE (ispB)"
                     /protein_id="CAA14934.1"
                     /db_xref="GI:3861034"
                     /translation="MNIIVKIQQNLKDEVTQLNDLIISCLKSDAELIEKVGKYLVEAG
                     GKRIRPLLTIITAKMFDYKGNNHIKLASAVEFIHAATLLHDDVVDNSTLRRFKPTANV
                     IWGSKTSILVGDFLFSQSFKLMVASGCIKAMNVLAKASVIISEGEVVQLVKLNERRII
                     TIDEYQQIVKSKTAELFGAACEVGAIIAEQVDRVSKDVQNFGRLLGTIFQVIDDLLDY
                     LGSDKQVGKNIGDDFLEGKVTLPLIFLYHKLEQDKQLWLENMLKSDKRTKDDFVKIRD
                     LMLKHAIYNETVNYLSSLENEANNLLNKIPVQNIYKYYLFSIIRFILYRSY"
     gene            1065. .1138
                     /gene="tRNA Arg (CCG)"
     tRNA            1065. .1138
                     /gene="tRNA Arg (CCG)"
     gene            complement(1194. .1265)
                     /gene="tRNA Gln (TTG)"
     tRNA            complement(1194. .1265)
                     /gene="tRNA Gln (TTG)"
     gene            complement(1316. .3091)
                     /gene="RP482"
     CDS             complement(1316. .3091)
                     /gene="RP482"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="CAA14935.1"
                     /db_xref="GI:3861035"
                     /translation="MRHDRINLLRKLFIEYNIEGYIIPSNDKYMNEYVPEYAKRLEYI
                     TGFTGSSGIAIICKDAAFFFTDGRYLEQANKELDLAFYKIYDLKEIYKTLDKNIKIGY
                     DPQLFTYQVLANLNINFHKINENLVDKIWYNKPLEPNSKVYLHDIKFAGVSHNDKIRK
                     CRETILSSSSVTVGCNKNNDDILVILDSASICWLLNLRASDVNYTPLMFAKVILTSTK
                     LYLFINPIRIDTEIINARPEITILPEKEFENILRDSKNRYLIDDSITSVHIMDLIANK
                     KVKKIVEPCLLAKACKNDIEIKHAIDFHIKDAVALCEFFAEFFLYHSSENVNSCFHSH
                     EIITEHSLCLKLTAQRAKQEGYVSDSFHAICGFQENSAIIHYRANPKTAKKIEGHGIL
                     LIDSGAQYKGATTDITRTIIVGIPTCEQKKRYTQVLKGHIALTRAKFPKNIVTGANLD
                     ILARQYLWQDMIDYPHGTGHGVGSFLSVHEGPQSINLSNKIILKAGMILSNEPGFYIP
                     GKYGIRIENLIYVKENNGWLEFETLSLVPYASKLIDVALLNIDEINYIKEYYNKIRAK
                     IYNLLSTQAKYWLNYEINLFLQSLL"
     gene            complement(3174. .4463)
                     /gene="RP483"
     CDS             complement(3174. .4463)
                     /gene="RP483"
                     /codon_start=1
                     /transl_table=11
                     /product="PUTRESCINE-ORNITHINE ANTIPORTER (potE)"
                     /protein_id="CAA14936.1"
                     /db_xref="GI:3861036"
                     /translation="MSQKLSFWAVFALVTGSQIGTSVFILPLSLAPFGVYSIWGWILS
                     LFGAMSIALVFSCLCTKFPKTGGPHVYVRASFGDTIAFFTGWTYWIISFVSTSIVVIS
                     AIGYLTPFFKSQTILDLILQLILLAAIAILNLKGPKIAGKVEFYLTLLKFVPLLVVGL
                     AALFHFNIDNIVIAKEVENFTIPTIMGRVALLTFWGFIGIECATTTAGTVKDPAKTIP
                     RAIIIGTCCVAFLYIINSIGIIGLIPASELINSKAPYADAATLLFGGTWSKVITVIAS
                     VICIGTLNAWVLTSGQIALGLAEDGLLPKFFAKKNSNNAPTYGIIISCLGITPLLLFT
                     SNNNFAKQITQIIDFSVIAFLFVYLICSLAFLKVIFSSKENFSYYYLFVAIISIIFCT
                     WVIYKTPFETLIIACSFTILGIPVYYGWYRLMNRIKL"
     gene            complement(4633. .4965)
                     /gene="RP484"
     CDS             complement(4633. .4965)
                     /gene="RP484"
                     /codon_start=1
                     /transl_table=11
                     /product="HESB PROTEIN (hesB2)"
                     /protein_id="CAA14937.1"
                     /db_xref="GI:3861037"
                     /translation="MKNVISLTDAAAKQVKLLIEKRAKPTFGIRVGIKSGGCAGQTYY
                     VEYADNKNQFDEVVEEKGVRILIDPKTLMYILGSEMDYVETNFKSQFTFTNPNEKANC
                     GCGKSFSV"
     gene            complement(4969. .5364)
                     /gene="RP485"
     CDS             complement(4969. .5364)
                     /gene="RP485"
                     /codon_start=1
                     /transl_table=11
                     /product="NIFU PROTEIN (nifU)"
                     /protein_id="CAA14938.1"
                     /db_xref="GI:3861038"
                     /translation="MAYSKKVIDHYENPRNVGSLDKKKKNVGTGLVGAPACGDVMKLQ
                     IEVGDDEIITDAKFKTFGCGSAIASSSLVTEWIKGKSVEDAKEIKNTEIAKELSLPPV
                     KLHCSLLAEDAIKAAIADYKQKRENKKDS"
     gene            complement(5430. .6662)
                     /gene="RP486"
     CDS             complement(5430. .6662)
                     /gene="RP486"
                     /codon_start=1
                     /transl_table=11
                     /product="NIFS PROTEIN HOMOLOG (spl1)"
                     /protein_id="CAA14939.1"
                     /db_xref="GI:3861039"
                     /translation="MNQQLKNLTLPIYMDYQSTTPIDPRVMEAMLPYFTTKFGNPHSR
                     SHSFGWEAENAVENARSMVAKVIGADSKEIIFTSGATESNNLVIKGIAKFYGNKKKHI
                     ITLVSEHKCVLNACRHLEQEGIKITYLPIKSNGIIDLETLKNAITDQTLLVSVMAVNN
                     EIGVIQPLKEIGKICRERNVFFHSDIAQGFGKIPINVNECNIDLASISGHKIYGPKGI
                     GALYIRKKPRVRVTPLINGGGQERGMRSGTLPTPLIVGLGIASEIAYNEMEKDTQHVN
                     YLFDRFLNNIHSKISEVYLNGDKDQRYKGNLNLSFAGVEGESIILAIKDLAVSSGSAC
                     TSASLEPSYVLRSIGISEELAHTSIRFGIGRFTTEQEIDYAVNLVCSKIDKLRRLSPL
```

```
                WEMMQEGVDLKKIRWTAH"
  gene          complement(6740. .7870)
                /gene="RP487"
  CDS           complement(6740. .7870)
                /gene="RP487"
                /codon_start=1
                /transl_table=11
                /product="NIFS PROTEIN HOMOLOG (spl1)"
                /protein_id="CAA14940.1"
                /db_xref="GI:3861040"
                /translation="MIYLDHNATTFIDPRVKEYIISLMDKELNPSSAHTSGRFAKNII
                ETARSQIATSLGITMSREYNITFTSSGTESNNLIMKNFYDGDIFISAIEHLSIYNHIN
                YAPNIKIIRVNKQGLVDLEHLEDLLSQSNASKKLVSVMIANNENGVLQDIAAISKITK
                KYNAKFHSDLVQSFGKISLNIKELGLDFATISGHKIGGGHGCGALISNSNFQLIPIII
                GGGQEKSVRSGTENVLAIAGFGLAAEFRRKDISKNYIKIRYLQERLEQKLKEYSNVNI
                ISNNVARLANTTLITVDNTDAQVKLIGFDLRNICVSSGSACSSGKISKSHVLTNMGIR
                EEEARSSIRISLSHTNTISDIKAFIKAFEEIYEYKSSCNHCI"
  gene          complement(7931. .8125)
                /gene="RP488"
  CDS           complement(7931. .8125)
                /gene="RP488"
                /codon_start=1
                /transl_table=11
                /product="unknown"
                /protein_id="CAA14941.1"
                /db_xref="GI:3861041"
                /translation="MILKIKEIYTVMAILELATNISNGPIILDKVSIKHNISLNYLAQ
                IFSKLKKSDLVKSIIESKDC"
  gene          complement(8277. .9143)
                /gene="RP489"
  CDS           complement(8277. .9143)
                /gene="RP489"
                /codon_start=1
                /transl_table=11
                /product="unknown"
                /protein_id="CAA14942.1"
                /db_xref="GI:3861042"
                /translation="MYLLSFILGGLNTLLRQLRLSISSIDFYKDVYKNYQGYGIKYLF
                TLSFIPSIIYCIFILNYIITLKDYFNGTSSSKVTDNIEYIINQLPKIKYNNSKISVEE
                VEPIYLYSKNNNKIFVIDTKNQVSNQEKSKVPFVLEENKLKINLVSNTKKHFPSIVNY
                SEIFKQNEVILTPEIIKKYFADNLLYAPNLFIFFGTPVIILFWFVTFLLERSIIVLLV
                YGLANLLTTKTSIQTSIRLVMFSSGIPIILQPVIIILIPELSILIQLLQMFTTFLVFV
                AILQINKSLSHI"
  gene          complement(9681. .10349)
                /gene="RP490"
  CDS           complement(9681. .10349)
                /gene="RP490"
                /codon_start=1
                /transl_table=11
                /product="unknown"
                /protein_id="CAA14943.1"
                /db_xref="GI:3861043"
                /translation="MKKLYHNPICPLSRQARFFLKELDIEFTTIKEDYWQFNKDFLKI
                NPAGTLPILEESYGLFVVGIYPFVEYLNNKYPNFNFLDTDIDICCEIRRLFFWFNNKF
                YFEVTKIIIDEKVIRSISKMSSPRTEFLRAAKNNLNYHLEYITSLLEKRSYLVSEFLT
                IADIAASCHLSVLDYFGEIYWDKWALVKHWYSLIKSRPAFRLLLQDRIPGFIPPSYYT
                DLDF"
  gene          complement(10535. .10893)
                /gene="RP491"
  misc_RNA      complement(10535. .10893)
                /gene="RP491"
  gene          11150. .13792
                /gene="RP492"
  CDS           11150. .13792
                /gene="RP492"
                /codon_start=1
                /transl_table=11
                /product="PYRUVATE,PHOSPHATE DIKINASE PRECURSOR (ppdK)"
                /protein_id="CAA14944.1"
                /db_xref="GI:3861044"

Alignment Scores:
Pred. No.:              6.01e-36        Length:         279110
```

```
Score:                  518.00          Matches:        99
Percent Similarity:     77.40%          Conservative:   38
Best Local Similarity:  55.93%          Mismatches:     40
Query Match:            58.53%          Indels:         0
DB:                     1               Gaps:           0

US-09-770-509-2 (1-178) x RPXX03 (1-279110)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              :::   |||:::|||       |||:::|||||||||||||||   |||:::|||
Db     230168 GCTAATCTACAAGGTGCTAATTTTGTAGTAGCGAATACTGATGCACAATCCCTTGAGCAT 230227

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              ||||||      :::||||||   ||||||         |||:::|||||||||||||||||
Db     230228 TCTTTATGCATTAACAAAATACAATTAGGCGTTTCTACTACTAGAGGTCTTGGTGCAGGT 230287

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   ||||||:::|||      :::|||::::::   :::   :::|||:::
Db     230288 GCATCACCTGAGGTTGGCGCGCTTGCTGCTCAAGAATCAGAAAATGAAATTCGTAGTTCA 230347

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||::::::|||||||||:::|||||||||   ||||||||||||||||||   ||||||
Db     230348 CTAGAAAATAGCAATATGGTATTTATTACAGCAGGTATGGGAGGTGGTACAGGTACTGGG 230407

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              :::||||||:::::|||   :::|||:::||||||||||||||||||||||||||||||
Db     230408 TCTGCACCTATTATTGCACGCATTGCGAAAGAACTAGGCATCCTTACGGTTGGGGTAGTG 230467

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||   :::|||   :::   |||:::   |||:::
Db     230468 ACTAAACCTTTCCACTTTGAAGGAGGTCATCGTATGAAAACAGCTGATAAAGGACTTATT 230527

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              ||||||   ::::::|||||||||||||||||:::|||||||||||||||      :::
Db     230528 GAACTACAGCAATTCGTTGATACTTTAATTGTAATACCGAACCAAAATCTATTCCGTATC 230587

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              |||::::::   ||||||      :::|||||:::   ||||||||||||||||
Db     230588 GCTAACGAACAAACAACATTTGCTGATGCTTTCAAAATGGCAGATGACGTATTACATGCA 230647

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||||||||||||||:::::   ||||||||||||||||||
Db     230648 GGTGTTAGAGGAGTAACAGATTTAATGATTATGCCTGGACTGATTAATCTT 230698
```

```
RESULT 11
WSU37260
LOCUS       WSU37260                697 bp    DNA     linear   BCT 29-AUG-1996
DEFINITION  Wolbachia sp. ftsZ gene, partial cds.
ACCESSION   U37260
VERSION     U37260.1  GI:1514677
KEYWORDS    .
SOURCE      Wolbachia sp.
  ORGANISM  Wolbachia sp.
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Wolbachieae; Wolbachia.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Tsagkarakou,A., Guillemaud,T., Rousset,F. and Navajas,M.
  TITLE     Molecular identification of a Wolbachia endosymbiont in a
            Tetranychus urticae strain (Acari: Tetranychidae)
  JOURNAL   Insect Mol. Biol. 5 (3), 217-221 (1996)
  MEDLINE   96392953
   PUBMED   8799741
REFERENCE   2  (bases 1 to 697)
  AUTHORS   Navajas,M., Tsagkarakou,A., Guillemaud,T. and Rousset,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-1995) Maria Navajas, Zoologie, I.N.R.A., 2, Place
            P. Viala, Montpellier, 34060, France
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="Wolbachia sp."
                     /specific_host="Tetranychus urticae"
                     /db_xref="taxon:956"
     gene            1. .697
```

```
                     /gene="ftsZ"
     CDS             <1. .>697
                     /gene="ftsZ"
                     /codon_start=3
                     /transl_table=11
                     /protein_id="AAC44314.1"
                     /db_xref="GI:1514678"
                     /translation="AVNNMIQSNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGL
                     GAGALPDVGKGAAEESIDEIMEHIKDSHMLFITAGMGGGTGTGAAPVIAKAAREARAV
                     VKDKGAKEKKILTVGVVTKPFGFEGVRRMRIAELGLEELQKYVDTLIVIPNQNLFRIA
                     NEKTTXVDAFQLADNVLHIGIRGVTDLMIMPGLINLDFADIETVMSEMGKAMIGTGEA
                     EGEDRAISAAEAA"
BASE COUNT      228 a    104 c    180 g    183 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:                7.62e-39        Length:         697
Score:                    517.00          Matches:        106
Percent Similarity:       72.49%          Conservative:   31
Best Local Similarity:    56.08%          Mismatches:     40
Query Match:              58.42%          Indels:         12
DB:                       1               Gaps:           1

US-09-770-509-2 (1-178) x WSU37260 (1-697)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::||||||   |||:::|||||||||||||||||   |||||||||   :::
Db         24 TCCAATTTGCAAGGGGTAAATTTTGTTGTAGCAAATACCGATGCTCAAGCGTTAGAGAAG 83

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              ||||||       ||||||   ||||||   ::::::|||||||||||||||||||||
Db         84 TCGTTATGTGATAAAAAAATTCAACTTGGTATTAACTTAACTAAGGGCCTTGGTGCCGGC 143

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   |||:::::||||||   :::|||||||:::         :::|||
Db        144 GCTTTGCCTGATGTTGGTAAAGGTGCAGCAGAAGAATCAATTGATGAGATTATGGAGCAT 203

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              ::::::|||||||:::|||||||||||||||   ||||||||||||||||||   ||||||
Db        204 ATAAAAGATAGCCATATGCTCTTTATCACAGCAGGGATGGGTGGTGGTACTGGAACAGGT 263

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeu---------------------- 93
              ||||||||||||:::|||      |||||||||
Db        264 GCTGCACCGGTAATTGCAAAAGCAGCCAGAGAAGCAAGAGCCGTAGTTAAAGATAAAGGA 323

Qy         94 -----------GlyIleLeuThrValGlyValValSerThrProPheArgSerGluGly 109
                         |||||||||||||||||||||:::   ||||||      ||||||
Db        324 GCAAAAGAAAAAAAGATACTGACTGTTGGAGTTGTAACTAAGCCGTTCGGTTTTGAAGGT 383

Qy        110 ProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyrValAspThr 129
              |||   |||:::|||      |||:::::||||||   ||||||||||||||||
Db        384 GTGCGACGTATGCGCATTGCAGAGCTTGGACTTGAAGAGTTGCAAAAATACGTAGATACA 443

Qy        130 LeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThrThrMetLeu 149
              |||||||||:::||||||||||||||||      :::|||:::::   ||||||   :::
Db        444 CTTATTGTCATTCCCAATCAAAATTTATTTAGAATTGCTAACGAGAAAACTACATTNGTN 503

Qy        150 GluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyValThrAspLeu 169
              :::|||||:::   ||||||:::||||||      |||:::::||||||||||||||||
Db        504 GACGCATTTCAACTCGCCGATAATGTTCTGCATATTGGCATAAGAGGAGTAACTGATTTG 563

Qy        170 IleValArgProGlyLeuIleAsnLeu 178
              ::::::   ||||||||||||||||||
Db        564 ATGATCATGCCAGGACTGATTAACCTT 590

RESULT 12
WSU95751
LOCUS       WSU95751                 868 bp    DNA     linear   BCT 11-AUG-1999
DEFINITION  Wolbachia sp. 123B cell division protein FtsZ (ftsZ) gene, partial
            cds.
ACCESSION   U95751
VERSION     U95751.1  GI:2078546
KEYWORDS    .
SOURCE      Wolbachia sp. 123B.
  ORGANISM  Wolbachia sp. 123B
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Wolbachieae; Wolbachia.
REFERENCE   1  (bases 1 to 868)
  AUTHORS   Jager,C.R., Pintureau,B. and Heddi,A.
  TITLE     Comparison between phylogenetic trees of some Trichogramma species
            and their Wolbachia endosymbionts
  JOURNAL   Russ. Entomol. J. 7 (3-4), 163-168 (1998)
REFERENCE   2  (bases 1 to 868)
  AUTHORS   Jager,C.R., Pintureau,B. and Heddi,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-1997) Laboratoire de Biologie Appliquee,
            INSA-Lyon, 20, av A. Einstein, Villeurbanne 69621, France
FEATURES             Location/Qualifiers
     source          1. .868
                     /organism="Wolbachia sp. 123B"
                     /strain="123B"
                     /specific_host="Trichogramma bourarachae"
                     /db_xref="taxon:59580"
     gene            <1. .>868
                     /gene="ftsZ"
     CDS             <1. .>868
                     /gene="ftsZ"
                     /codon_start=2
                     /transl_table=11
                     /product="cell division protein FtsZ"
                     /protein_id="AAB54069.1"
                     /db_xref="GI:2078547"
                     /translation="QSNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALP
                     DVGKGAAEESIDEIMEHIKDSHMLFITAGMGGGTGTGAAPVIAKAAREARAAVKDRAP
                     KEKKILTVGVVTKPFGFEGVRRMRIAELGLEELQKYVDTLIVIPNQNLFRIANEKTTF
                     SDAFKLADNVLHIGIRGVTDLMVMPGLINLDFADIETVMSEMGKAMIGTGEAEGEDRA
                     ISAAEAAISNPLLDNVSMKGAQGILINITGGGDMTLFEVDAAANRVREEVDENANIIF
                     GATFDQAMEGRVR"
BASE COUNT      283 a    133 c    224 g    228 t
ORIGIN

Alignment Scores:
Pred. No.:                9.81e-39        Length:         868
Score:                    517.00          Matches:        107
Percent Similarity:       71.96%          Conservative:   29
Best Local Similarity:    56.61%          Mismatches:     41
Query Match:              58.42%          Indels:         12
DB:                       1               Gaps:           1

US-09-770-509-2 (1-178) x WSU95751 (1-868)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::||||||   |||:::|||||||||||||||||   |||||||||   :::
Db          5 TCTAATTTGCAAGGAGTGAATTTTGTTGTCGCAAATACCGATGCTCAAGCACTAGAAAAG 64

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              ||||||       ||||||   ||||||   ::::::|||||||||||||||||||||
Db         65 TCGTTATGTGATAAAAAAATTCAGCTTGGCATTAATTTAACTAAGGGGCTTGGTGCTGGT 124

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   |||:::::||||||   :::|||||||:::         :::|||
Db        125 GCTTTGCCTGATGTTGGCAAAGGTGCAGCAGAAGAGTCAATTGATGAGATTATGGAGCAT 184

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              ::::::|||||||:::|||||||||||||||   ||||||||||||||||||   ||||||
Db        185 ATAAAAGATAGTCATATGCTTTTCATCACAGCAGGAATGGGCGGTGGTACTGGAACCGGT 244

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeu---------------------- 93
              ||||||||||||:::|||      |||||||||
Db        245 GCAGCACCGGTAATTGCAAAAGCAGCCAGAGAAGCAAGAGCCGCAGTTAAGGATAGAGCG 304

Qy         94 -----------GlyIleLeuThrValGlyValValSerThrProPheArgSerGluGly 109
                         |||||||||||||||||||||:::   ||||||      ||||||
Db        305 CCAAAAGAAAAAAAGATATTGACTGTTGGAGTTGTAACTAAACCGTTCGGTTTTGAAGGT 364

Qy        110 ProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyrValAspThr 129
```

```
                 |||   |||:::|||      |||::::::|||||   ||||||||||||||||
Db        365 GTGCGCCGTATGCGCATTGCAGAGCTTGGACTTGAAGAACTGCAAAAATACGTGGATACA 424

Qy        130 LeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThrThrMetLeu 149
              |||||||||:::|||||||||||||||      :::|||::::::   ||||||
Db        425 CTTATTGTCATTCCAAATCAGAATTTATTTAGAATTGCAAATGAAAAAACTACATTTTCT 484

Qy        150 GluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyValThrAspLeu 169
              :::|||||:::   ||||||:::||||||      |||:::::::|||||||||||||||
Db        485 GATGCATTTAAACTTGCTGATAATGTTCTGCACATTGGCATCAGAGGAGTAACTGACTTG 544

Qy        170 IleValArgProGlyLeuIleAsnLeu 178
              :::|||   ||||||||||||||||||
Db        545 ATGGTCATGCCAGGGCTTATCAATCTT 571

RESULT 13
WSU95753
LOCUS       WSU95753                868 bp    DNA     linear   BCT 11-AUG-1999
DEFINITION  Wolbachia sp. t191 cell division protein FtsZ (ftsZ) gene, partial
            cds.
ACCESSION   U95753
VERSION     U95753.1  GI:2078550
KEYWORDS    .
SOURCE      Wolbachia sp. t191.
  ORGANISM  Wolbachia sp. t191
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Wolbachieae; Wolbachia.
REFERENCE   1  (bases 1 to 868)
  AUTHORS   Jager,C.R., Pintureau,B. and Heddi,A.
  TITLE     Comparison between phylogenetic trees of some Trichogramma species
            and their Wolbachia endosymbionts
  JOURNAL   Russ. Entomol. J. 7 (3-4), 163-168 (1998)
REFERENCE   2  (bases 1 to 868)
  AUTHORS   Jager,C.R., Pintureau,B. and Heddi,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-1997) Laboratoire de Biologie Appliquee,
            INSA-Lyon, 20, av A. Einstein, Villeurbanne 69621, France
FEATURES             Location/Qualifiers
     source          1. .868
                     /organism="Wolbachia sp. t191"
                     /strain="t191"
                     /specific_host="Trichogramma pretiosum"
                     /db_xref="taxon:59582"
     gene            <1. .>868
                     /gene="ftsZ"
     CDS             <1. .>868
                     /gene="ftsZ"
                     /codon_start=2
                     /transl_table=11
                     /product="cell division protein FtsZ"
                     /protein_id="AAB54071.1"
                     /db_xref="GI:2078551"
                     /translation="QSNLQGVNFIVANTDAQALEKSLCDKKIQLGINLTKGLGAGALP
                     DIGKGAAEESIDEIMEHIRDSHMLFITAGMGGGTGTGAAPVIAKAAREARAVVKDKGA
                     KEKKILTVGVVTKPFGFEGVRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTF
                     ADAFQLADNVLHIGIRGVTDLMIMPGLINLDFADIETVMSEMGKAMIGTGEAEGEDRA
                     ISAAEAAISNPLLDNVSMKGAQGILINITGGGDMTLFEVDAAANRVREEVDENANIIF
                     GATFDQAMEGRVR"
BASE COUNT      282 a    123 c    229 g    234 t
ORIGIN

Alignment Scores:
Pred. No.:                1.5e-38        Length:         868
Score:                    515.00         Matches:        107
Percent Similarity:       71.43%         Conservative:   28
Best Local Similarity:    56.61%         Mismatches:     42
Query Match:              58.19%         Indels:         12
DB:                       1              Gaps:           1

US-09-770-509-2 (1-178) x WSU95753 (1-868)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::||||||   ||||||||||||||||||||||   |||||||||   :::

Db          5 TCCAATTTACAAGGAGTAAATTTTATTGTAGCAAATACCGATGCTCAAGCGTTAGAGAAG 64

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              ||||||          ||||||   ||||||   ::::::|||||||||||||||||||||
Db         65 TCATTATGTGATAAAAAAATTCAGCTTGGTATTAACTTAACTAAGGGCCTTGGCGCCGGC 124

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   |||::::::||||||   :::||||||:::         :::|||
Db        125 GCTTTGCCTGATATTGGTAAAGGTGCAGCAGAAGAATCAATTGATGAAATTATGGAGCAT 184

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              ::::::||||||:::|||||||||||||||   ||||||||||||||||   ||||||
Db        185 ATAAGGGACAGCCATATGCTCTTTATCACAGCAGGGATGGGTGGTGGTACTGGAACAGGT 244

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeu---------------------- 93
              ||||||||||||:::|||      ||||||||
Db        245 GCTGCACCGGTAATTGCAAAAGCAGCCAGAGAAGCAAGAGCGGTGGTTAAAGATAAAGGA 304

Qy         94 ------------GlyIleLeuThrValGlyValValSerThrProPheArgSerGluGly 109
                          ||||||||||||||||||||||:::   ||||||      ||||||
Db        305 GCAAAAGAAAAAAAGATACTGACTGTTGGAGTTGTAACTAAGCCGTTCGGTTTTGAGGGT 364

Qy        110 ProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyrValAspThr 129
              |||   |||   |||      |||:::::||||||   ||||||||||||||
Db        365 GTGCGACGTATGCGCACTGCAGAGCTTGGACTTGAAGAGTTGCAAAAATACGTAGATACA 424

Qy        130 LeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThrThrMetLeu 149
              |||||||||:::|||||||||||||||      :::|||::::::   ||||||
Db        425 CTTATTGTCATTCCCAATCAAAATTTATTTAGAATTGCTAACGAGAAAACTACATTTGCT 484

Qy        150 GluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyValThrAspLeu 169
              :::|||||:::   ||||||:::||||||      |||:::::::|||||||||||||||
Db        485 GATGCATTTCAACTCGCCGATAATGTTCTGCATATTGGCATAAGAGGAGTAACTGATTTG 544

Qy        170 IleValArgProGlyLeuIleAsnLeu 178
              ::::::   ||||||||||||||||||
Db        545 ATGATCATGCCAGGACTGATTAATCTT 571

RESULT 14
WSU95749
LOCUS       WSU95749                868 bp    DNA     linear   BCT 11-AUG-1999
DEFINITION  Wolbachia sp. 1032 cell division protein FtsZ (ftsZ) gene, partial
            cds.
ACCESSION   U95749
VERSION     U95749.1  GI:2078542
KEYWORDS    .
SOURCE      Wolbachia sp. 1032.
  ORGANISM  Wolbachia sp. 1032
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Wolbachieae; Wolbachia.
REFERENCE   1  (bases 1 to 868)
  AUTHORS   Jager,C.R., Pintureau,B. and Heddi,A.
  TITLE     Comparison between phylogenetic trees of some Trichogramma species
            and their Wolbachia endosymbionts
  JOURNAL   Russ. Entomol. J. 7 (3-4), 163-168 (1998)
REFERENCE   2  (bases 1 to 868)
  AUTHORS   Jager,C.R., Pintureau,B. and Heddi,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-1997) Laboratoire de Biologie Appliquee,
            INSA-Lyon, 20, av A. Einstein, Villeurbanne 69621, France
FEATURES             Location/Qualifiers
     source          1. .868
                     /organism="Wolbachia sp. 1032"
                     /strain="1032"
                     /specific_host="Trichogramma cordubensis"
                     /db_xref="taxon:59578"
     gene            <1. .>868
                     /gene="ftsZ"
     CDS             <1. .>868
                     /gene="ftsZ"
                     /codon_start=2
                     /transl_table=11
                     /product="cell division protein FtsZ"
```

```
                     /protein_id="AAB54067.1"
                     /db_xref="GI:2078543"
                     /translation="QSNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALP
                     DIGKGAAEESIDEIMEHIRDSHMLFITAGMGGGTGTGAAPVIAEAARERARAVVKDKGA
                     KEKKILTVGVVTKPFGFEGVRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTF
                     ADAFQLADNVLHIGIRGVTDLMIMPGLINLDFADIETVMSEMGKAMIGTGEAEGEDRA
                     ISAAEAAISNPLLDNVSMKGAQGILINITGGGDMTLFEVDAAANRVREEVDENANIIF
                     GATFDQVMEGRVR"
BASE COUNT      278 a    121 c    234 g    235 t
ORIGIN

Alignment Scores:
Pred. No.:                1.85e-38       Length:          868
Score:                    514.00         Matches:         106
Percent Similarity:       71.43%         Conservative:    29
Best Local Similarity:    56.08%         Mismatches:      42
Query Match:              58.08%         Indels:          12
DB:                       1              Gaps:            1

US-09-770-509-2 (1-178) x WSU95749 (1-868)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::|||||||   |||:::|||||||||||||||||   |||||||||   :::
Db          5 TCCAATTTACAAGGAGTAAATTTTGTTGTAGCAAATACCGATGCTCAGGCGTTAGAGAAG 64

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              ||||||         ||||||   ||||||   ::::::||||||||||||||||||||
Db         65 TCATTATGTGATAAAAAAATTCAGCTTGGTATTAACTTAACTAAGGGCCTTGGCGCCGGA 124

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   |||:::::|||||   :::|||||:::         :::|||
Db        125 GCTTTGCCTGATATTGGTAAAGGTGCAGCAGAAGAATCAATTGATGAAATTATGGAGCAT 184

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              ::::::|||||:::||||||||||||||   ||||||||||||||||   ||||||
Db        185 ATAAGGGACAGCCATATGCTCTTTATCACAGCAGGGATGGGTGGTGGTACTGGAACAGGT 244

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeu---------------------- 93
              ||||||||||||:::|||      |||||||||
Db        245 GCTGCACCGGTGATTGCAGAAGCAGCCAGAGAAGCAAGAGCGGTGGTTAAAGATAAAGGA 304

Qy         94 ------------GlyIleLeuThrValGlyValValSerThrProPheArgSerGluGly 109
                         ||||||||||||||||||||:::   ||||||      ||||||
Db        305 GCAAAAGAAAAAAAGATACTGACTGTTGGAGTTGTAACTAAGCCGTTCGGTTTTGAAGGT 364

Qy        110 ProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyrValAspThr 129
              |||   |||   |||      |||:::::||||||   ||||||||||||||
Db        365 GTGCGACGTATGCGCACTGCAGAGCTTGGACTTGAAGAGTTGCAAAAATACGTAGATACA 424

Qy        130 LeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThrThrMetLeu 149
              |||||||||:::||||||||||||||||      :::|||::::::   ||||||
Db        425 CTTATTGTCATTCCCAATCAAAATTTATTTAGAATTGCTAACGAGAAAACTACATTTGCT 484

Qy        150 GluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyValThrAspLeu 169
              :::|||||:::   ||||||::||||||      |||:::::|||||||||||||||||
Db        485 GATGCATTTCAACTCGCCGATAATGTTCTGCATATTGGCATAAGAGGAGTAACTGATTTG 544

Qy        170 IleValArgProGlyLeuIleAsnLeu 178
              ::::::   ||||||||||||||||||
Db        545 ATGATCATGCCAGGACTGATTAATCTT 571

RESULT 15
WSU95750
LOCUS       WSU95750                868 bp    DNA     linear   BCT 11-AUG-1999
DEFINITION  Wolbachia sp. 1148 cell division protein FtsZ (ftsZ) gene, partial
            cds.
ACCESSION   U95750
VERSION     U95750.1  GI:2078544
KEYWORDS    .
SOURCE      Wolbachia sp. 1148.
  ORGANISM  Wolbachia sp. 1148
            Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
            Rickettsiaceae; Wolbachieae; Wolbachia.
REFERENCE   1  (bases 1 to 868)
  AUTHORS   Jager,C.R., Pintureau,B. and Heddi,A.
  TITLE     Comparison between phylogenetic trees of some Trichogramma species
            and their Wolbachia endosymbionts
  JOURNAL   Russ. Entomol. J. 7 (3-4), 163-168 (1998)
REFERENCE   2  (bases 1 to 868)
  AUTHORS   Jager,C.R., Pintureau,B. and Heddi,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-1997) Laboratoire de Biologie Appliquee,
            INSA-Lyon, 20, av A. Einstein, Villeurbanne 69621, France
FEATURES             Location/Qualifiers
     source          1. .868
                     /organism="Wolbachia sp. 1148"
                     /strain="1148"
                     /specific_host="Trichogramma cordubensis"
                     /db_xref="taxon:59579"
     gene            <1. .>868
                     /gene="ftsZ"
     CDS             <1. .>868
                     /gene="ftsZ"
                     /codon_start=2
                     /transl_table=11
                     /product="cell division protein FtsZ"
                     /protein_id="AAB54068.1"
                     /db_xref="GI:2078545"
                     /translation="QSNLQGVNFVVANTDAQALEKSLCDKKIQLGINLTKGLGAGALP
                     DIGKGAAEESIDEIMEHIRDSHMLFITAGMGGGTGTGAAPVIAKAARERARAVVKDKGA
                     KEKKILTVGVVTKPFGFEGVRRMRTAELGLEELQKYVDTLIVIPNQNLFRIANEKTTF
                     ADAFQLADNVLHIGIRGVTDLMIMPGLINLDFADIETVMSEMGKAMIGTGEAEGEDRA
                     ISAAEAAISNPLLDNVSMKGAQGILINITGGGDMTLFEVDAAANRVREEVDENANIIF
                     GATFDQVMEGRVR"
BASE COUNT      279 a    121 c    233 g    235 t
ORIGIN

Alignment Scores:
Pred. No.:                1.85e-38       Length:          868
Score:                    514.00         Matches:         106
Percent Similarity:       71.43%         Conservative:    29
Best Local Similarity:    56.08%         Mismatches:      42
Query Match:              58.08%         Indels:          12
DB:                       1              Gaps:            1

US-09-770-509-2 (1-178) x WSU95750 (1-868)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::|||||||   |||:::|||||||||||||||||   |||||||||   :::
Db          5 TCCAATTTACAAGGAGTAAATTTTGTTGTAGCAAATACCGATGCTCAGGCGTTAGAGAAG 64

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              ||||||         ||||||   ||||||   ::::::||||||||||||||||||||
Db         65 TCATTATGTGATAAAAAAATTCAGCTTGGTATTAACTTAACTAAGGGCCTTGGCGCCGGA 124

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   |||:::::|||||   :::|||||:::         :::|||
Db        125 GCTTTGCCTGATATTGGTAAAGGTGCAGCAGAAGAATCAATTGATGAAATTATGGAGCAT 184

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              ::::::|||||:::||||||||||||||   ||||||||||||||||   ||||||
Db        185 ATAAGGGACAGCCATATGCTCTTTATCACAGCAGGGATGGGTGGTGGTACTGGAACAGGT 244

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeu---------------------- 93
              ||||||||||||:::|||      |||||||||
Db        245 GCTGCACCGGTGATTGCAAAAGCAGCCAGAGAAGCAAGAGCGGTGGTTAAAGATAAAGGA 304

Qy         94 ------------GlyIleLeuThrValGlyValValSerThrProPheArgSerGluGly 109
                         ||||||||||||||||||||:::   ||||||      ||||||
Db        305 GCAAAAGAAAAAAAGATACTGACTGTTGGAGTTGTAACTAAGCCGTTCGGTTTTGAAGGT 364

Qy        110 ProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyrValAspThr 129
              |||   |||   |||      |||:::::||||||   ||||||||||||||
Db        365 GTGCGACGTATGCGCACTGCAGAGCTTGGACTTGAAGAGTTGCAAAAATACGTAGATACA 424

Qy        130 LeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThrThrMetLeu 149
```

```
            |||||||||:::|||||||||||||||      :::|||::::::   ||||||
Db        425 CTTATTGTCATTCCCAATCAAAATTTATTTAGAATTGCTAACGAGAAAACTACATTTGCT 484

Qy        150 GluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyValThrAspLeu 169
            :::|||||||:::   ||||||:::||||||      |||::::::|||||||||||||||
Db        485 GATGCATTTCAACTCGCCGATAATGTTCTGCATATTGGCATAAGAGGAGTAACTGATTTG 544

Qy        170 IleValArgProGlyLeuIleAsnLeu 178
            ::::::   ||||||||||||||||||
Db        545 ATGATCATGCCAGGACTGATTAATCTT 571


Search completed: June  2, 2003, 09:39:19
Job time : 2508 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:44:29 ; Search time 245 Seconds
(without alignments)
1636.146 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09770509/runat_27052003_083647_18810/app_query.fasta_1.327
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09770509_@CGN_1_1_263_@runat_27052003_083647_18810 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		510	57.6	213251	24	ABQ67193	Listeria innocua c
2	c	510	57.6	2944528	24	ABA03041	Listeria monocytog
3	c	510	57.6	3011208	24	ABQ69245	Listeria innocua D
4		509	57.5	950	24	ABK77373	Bacillus clausii g
5		506	57.2	111309	20	AAX20250	Borrelia burgdorfe
6		506	57.2	910715	20	AAX20248	Borrelia burgdorfe
7		502	56.7	1134	24	ABK72889	Bacillus lichenifo
8		495	55.9	1233	23	AAS52905	Enterococcus faeca
9		495	55.9	1239	23	AAS51272	Enterococcus faeca
10	c	495	55.9	32768	20	AAX13065	Enterococcus faeca
11	c	493	55.7	2365589	24	ABA90521	Genomic sequence o
12		485	54.8	1140	22	AAH52021	Mycobacterium tube
13	c	485	54.8	4403765	22	AAI99683	Mycobacterium tube
14	c	485	54.8	4411529	22	AAI99682	Mycobacterium tube
15		480	54.2	1172	23	AAS52133	Staphylococcus aur
16		480	54.2	1173	21	AAA39395	Staphylococcus aur
17		480	54.2	1173	21	AAZ49958	S. pneumoniae mono
18		480	54.2	1179	23	AAS54988	Staphylococcus aur
19		480	54.2	1185	24	ABN92534	Staphylococcus epi
20		472	53.3	1008	19	AAV65296	DNA of filamentous
21		472	53.3	1260	20	AAX19269	Streptococcus pneu
22		472	53.3	1260	21	AAZ49959	Monomeric filament
23		472	53.3	1260	23	AAS55744	Streptococcus pneu
24		472	53.3	2702	19	AAV65251	DNA of filamentous
25		472	53.3	21706	19	AAV52169	Streptococcus pneu
26		470	53.1	1326	22	AAH67325	C glutamicum codin
27		470	53.1	3931	18	AAV74653	Staphylococcus aur
28	c	470	53.1	349980	22	AAH68531	C glutamicum codin
29		467	52.8	1278	24	ABN66155	Streptococcus poly
30		467	52.8	2155561	24	ABN71527	Streptococcus poly
31	c	466	52.7	29112	23	AAS59609	Propionibacterium
32		464	52.4	1185	23	AAS54272	Pseudomonas aerugi
33		464	52.4	1317	24	ABN66156	Streptococcus poly
34		464	52.4	7143	19	AAV58287	S. pyogenes SP-26-
35		463	52.3	1611	21	AAA47312	FtsZ2 gene sequenc
36		463	52.3	1611	21	AAA47339	Type II FtsZ seque
37		463	52.3	1628	19	AAV05550	Arabidopsis chloro
38	c	461	52.1	640681	24	ABA92787	Buchnera sp. genom
39		452	51.1	1152	21	AAA65913	E. coli proliferat
40		452	51.1	1152	21	AAA39397	Escherichia coli m
41		452	51.1	1152	23	AAS52299	E. coli DNA for ce
42		452	51.1	1255	21	AAA47315	FtsZ1 gene sequenc
43		447	50.5	1152	24	ABK11439	DNA encoding E. co
44		447	50.5	1266	23	AAS53458	Haemophilus influe
45		447	50.5	1830121	17	AAT42063	Haemophilus influe

ALIGNMENTS

RESULT 1
ABQ67193
ID ABQ67193 standard; DNA; 213251 BP.
XX
AC ABQ67193;
XX
DT 29-AUG-2002 (first entry)
XX
DE Listeria innocua contig DNA sequence #6.
XX
KW Antibacterial; Listeria; food contamination; mutational analysis;
KW infection; ds.
XX
OS Listeria innocua.
XX
PN WO200228891-A2.
XX
PD 11-APR-2002.

```
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F,  Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators  -
XX
PS   Claim 5; SEQ ID 6; 180pp; French.
XX
CC   The present invention relates to nucleic acid sequences
CC   (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC   and primers for identification and/or detection of Listeria (e.g. as
CC   contaminants in foods, or mutational analysis) and for analysis of
CC   gene expression. Proteins encoded by the nucleic acid sequences can be
CC   used to screen for compounds that modulate gene expression, replication
CC   and pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in
CC   anti-Listeria vaccines.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 213251 BP; 70939 A; 36026 C; 43257 G; 63029 T; 0 other;

Alignment Scores:
Pred. No.:                1.34e-43      Length:          213251
Score:                    510.00        Matches:         98
Percent Similarity:       76.00%        Conservative:    35
Best Local Similarity:    56.00%        Mismatches:      42
Query Match:              57.63%        Indels:          0
DB:                       24            Gaps:            0

US-09-770-509-2 (1-178) x ABQ67193 (1-213251)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||      |||||||||   |||||||||      :::
Db     170157 GTTCAAGGAGTAGAATTTATCTCCGTTAATACAGACGCTCAAGCACTTAATTTAGCAAAA 170216

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||      :::|||:::||||||||||||||:::
Db     170217 GCAGAAACAAAATTACAAATCGGTACAAAATTAACGCGTGGTTTAGGCGCGGGTGCTGTA 170276

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||:::::::|||||:::   :::      |||:::      |||:::
Db     170277 CCTGAAATTGGTAAAAAAGCAGCAGAAGAAAGTCGCGAACAAATTGAAGAAGCTTTAAAA 170336

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||:::|||:::|||:::|||   ||||||||||||||||||   ||||||||||||
Db     170337 GGCTCTGATATGGTATTCGTAACTGCTGGGATGGGCGGCGGAACTGGAACTGGTGCTGCA 170396

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   :::|||:::|||:::|||   ||||||||||||||||||:::
Db     170397 CCTGTTATCGCTCAAATCGCAAAAGAAATGGGCGCTCTAACTGTCGGTGTTGTTACACGA 170456

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||||||:::   |||      |||   :::   :::
Db     170457 CCATTTGGTTTTGAAGGACCAAAACGTACGAAACAAGCCCTAACTGGAACAGAAGCAATG 170516

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||||||||:::||||||      ||||||   :::   |||
Db     170517 AAAGAAGCGGTGGATACTTTAATTGTTATCCCTAACGACCGTTTACTTCAAATTGTTGAT 170576

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   ||||||||||||||||||   ||||||:::||||||   :::|||||||
```

```
Db     170577 AAAAATACGCCAATGCTTGAAGCTTTCCGTGAAGCAGATAATGTTTTACGTCAAGGGGTA 170636

Qy         164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
               :::|||:::::::|||||||||      ||||||||||||||||||
Db     170637 CAAGGGATTTCTGATTTGATTGCCGTTCCTGGTTTAATTAACTTA 170681

RESULT 2
ABA03041/c
ID   ABA03041 standard; DNA; 2944528 BP.
XX
AC   ABA03041;
XX
DT   05-FEB-2002  (first entry)
XX
DE   Listeria monocytogenes EGD-e genome sequence.
XX
KW   Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW   vitamin B12; bacterial infection; disease; ds.
XX
OS   Listeria monocytogenes.
XX
PN   WO200177335-A2.
XX
PD   18-OCT-2001.
XX
PF   11-APR-2001; 2001WO-FR01118.
XX
PR   11-APR-2000; 2000FR-0004629.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Buchrieser C,  Frangeul L,  Couve E,  Rusniok C,  Fsihi H,  Dehoux P;
PI   Dussurget O,  Chetouani F,  Nedjari H,  Glaser P,  Kunst F,  Cossart P;
PI   Daniels J,  Goebel W,  Kreft J,  Kuhn M,  Ng E,  Vazquez-Boland JA;
PI   Dominguez-Bernal G,  Garrido-Garcia P,  Tierrez-Martinez A,  Amend A;
PI   Chakraborty T,  Domann E,  Hain T,  Berche P,  Charbit A,  Durant L;
PI   Perez-Diaz J,  Baquero F,  Garcia Del Portillo F,  Gomez-Lopez N;
PI   Maduenio E,  De Pablos B,  Wehland J,  Kaerst U,  Entian K,  Hauf J;
PI   Rose M,  Voss H;
XX
DR   WPI; 2002-010914/01.
XX
PT   Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT   and prevention of Listeria and related bacterial infections, and
PT   related polypeptides  -
XX
PS   Claim 1; SEQ ID No 1; 192pp; French.
XX
CC   The present sequence is the genome sequence of Listeria monocytogenes
CC   EGD-e. This sequence and fragments of this sequence are useful for
CC   selecting probes and primers for detecting genes in L. monocytogenes and
CC   related organisms, and to study genetic polymorphisms and other genomes.
CC   Proteins (ABB47297-ABB50149) expressed from the present sequence are
CC   useful for raising specific antibodies, identification of L.
CC   monocytogenes and related organisms, and for biosynthesis and
CC   biodegradation, especially biosynthesis of Vitamin B12. This sequence and
CC   proteins encoded by it are also useful for selecting compounds that
CC   regulate gene expression and cell replication and modulate L.
CC   monocytogenes-related diseases. In addition, this sequence and proteins
CC   encoded by it are useful in pharmaceutical and vaccines compositions for
CC   the treatment or prevention of infections by L. monocytogenes and related
CC   organisms.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other;

Alignment Scores:
Pred. No.:                4.17e-42      Length:          2944528
Score:                    510.00        Matches:         98
Percent Similarity:       76.00%        Conservative:    35
Best Local Similarity:    56.00%        Mismatches:      42
```

```
Query Match:            57.63%      Indels:        0
DB:                     24          Gaps:          0

US-09-770-509-2 (1-178) x ABA03041 (1-2944528)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||      |||||||||   |||||||||      :::
Db    2113565 GTTCAAGGAGTAGAATTTATCTCCGTTAATACAGACGCTCAAGCACTTAACTTAGCAAAA 2113506

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||      :::|||:::|||||||||||||||:::
Db    2113505 GCAGAAACAAAATTACAAATCGGTACAAAATTAACGCGTGGTTTAGGTGCGGGAGCTGTA 2113446

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||:::::::||||||:::   :::      |||:::      |||:::
Db    2113445 CCTGAAATTGGTAAAAAAGCTGCAGAAGAAAGCCGCGAACAAATTGAAGAAGCTTTAAAA 2113386

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||:::|||:::|||:::|||   ||||||||||||||||||   ||||||||||||
Db    2113385 GGCTCTGATATGGTATTCGTAACTGCTGGAATGGGCGGCGGAACTGGAACTGGGGCTGCA 2113326

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   :::|||:::|||:::|||   |||||||||||||||||:::
Db    2113325 CCTGTTATCGCTCAAATCGCAAAAGAAATGGGCGCTTTAACAGTAGGTGTTGTTACTCGA 2113266

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   ||||||:::   |||      |||   :::   :::
Db    2113265 CCATTTGGTTTTGAAGGACCAAAACGTACGAAACAAGCCTTAACTGGAACAGAAGCAATG 2113206

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::   ||||||||||||||||||:::||||||      ||||||   :::   |||
Db    2113205 AAAGAAGCGGTGGATACGTTAATTGTTATTCCTAATGACCGTTTACTTCAAATTGTGGAT 2113146

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   ||||||||||||||||   ||||||:::|||||   :::|||||||
Db    2113145 AAAAATACACCGATGCTTGAAGCTTTCCGTGAAGCAGATAATGTTTTACGTCAAGGTGTA 2113086

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::|||||||||      ||||||||||||||||||
Db    2113085 CAAGGGATTTCTGATTTGATTGCCGTTCCTGGTTTAATTAACTTA 2113041

RESULT 3
ABQ69245/c
ID   ABQ69245 standard; DNA; 3011208 BP.
XX
AC   ABQ69245;
XX
DT   29-AUG-2002  (first entry)
XX
DE   Listeria innocua DNA sequence #684.
XX
KW   Antibacterial; Listeria; food contamination; mutational analysis;
KW   infection; ds.
XX
OS   Listeria innocua.
XX
PN   WO200228891-A2.
XX
PD   11-APR-2002.
XX
PF   04-OCT-2001; 2001WO-FR03061.
XX
PR   04-OCT-2000; 2000FR-0012697.
XX
PA   (INSP ) INST PASTEUR.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Kunst F,  Glaser P;
XX
DR   WPI; 2002-332479/37.
XX
PT   New genomic sequences from Listeria species, useful for detection,
PT   treatment and prevention of infection, also related polypeptides,
PT   antibodies and modulators  -
XX
PS   Claim 5; SEQ ID 2058; 180pp; French.
XX
CC   The present invention relates to nucleic acid sequences
CC   (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes
CC   and primers for identification and/or detection of Listeria (e.g. as
CC   contaminants in foods, or mutational analysis) and for analysis of
CC   gene expression. Proteins encoded by the nucleic acid sequences can be
CC   used to screen for compounds that modulate gene expression, replication
CC   and pathogenicity of Listeria (potential therapeutic agents), also for
CC   treating infections by Listeria, and are useful as immunogens in
CC   anti-Listeria vaccines.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3011208 BP; 941651 A; 568176 C; 559189 G; 942192 T; 0 other;

Alignment Scores:
Pred. No.:              4.29e-42    Length:        3011208
Score:                  510.00      Matches:       98
Percent Similarity:     76.00%      Conservative:  35
Best Local Similarity:  56.00%      Mismatches:    42
Query Match:            57.63%      Indels:        0
DB:                     24          Gaps:          0

US-09-770-509-2 (1-178) x ABQ69245 (1-3011208)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||      |||||||||   |||||||||      :::
Db    2161845 GTTCAAGGAGTAGAATTTATCTCCGTTAATACAGACGCTCAAGCACTTAATTTAGCAAAA 2161786

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||      :::|||:::|||||||||||||||:::
Db    2161785 GCAGAAACAAAATTACAAATCGGTACAAAATTAACGCGTGGTTTAGGCGCGGGTGCTGTA 2161726

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||:::::::|||||||:::   :::      |||:::      |||:::
Db    2161725 CCTGAAATTGGTAAAAAAGCAGCAGAAGAAAGTCGCGAACAAATTGAAGAAGCTTTAAAA 2161666

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||:::|||:::|||:::|||   ||||||||||||||||||   ||||||||||||
Db    2161665 GGCTCTGATATGGTATTCGTAACTGCTGGGATGGGCGGCGGAACTGGAACTGGTGCTGCA 2161606

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   :::||||:::|||:::|||   |||||||||||||||||:::
Db    2161605 CCTGTTATCGCTCAAATCGCAAAAGAAATGGGCGCTCTAACTGTCGGTGTTGTTACACGA 2161546

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||||||:::   |||      |||   :::   :::
Db    2161545 CCATTTGGTTTTGAAGGACCAAAACGTACGAAACAAGCCCTAACTGGAACAGAAGCAATG 2161486

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::   ||||||||||||||||||:::||||||      ||||||   :::   |||
Db    2161485 AAAGAAGCGGTGGATACTTTAATTGTTATCCCTAACGACCGTTTACTTCAAATTGTTGAT 2161426

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   ||||||||||||||||   ||||||:::||||||   :::|||||||
Db    2161425 AAAAATACGCCAATGCTTGAAGCTTTCCGTGAAGCAGATAATGTTTTACGTCAAGGGGTA 2161366

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::|||||||||      ||||||||||||||||||
Db    2161365 CAAGGGATTTCTGATTTGATTGCCGTTCCTGGTTTAATTAACTTA 2161321

RESULT 4
ABK77373
ID   ABK77373 standard; DNA; 950 BP.
XX
AC   ABK77373;
XX
DT   13-AUG-2002  (first entry)
XX
```

DE Bacillus clausii genomic sequence tag (GST) #216.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus clausii.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US31437.
XX
PR 06-OCT-2000; 2000US-0680598.
PR 27-MAR-2001; 2001US-279526P.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
PA (NOVO) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array -
XX
PS Claim 11; SEQ ID NO 4664; 200pp; English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive
CC follow-up characterisation is unnecessary, when one spot on an array
CC equals one gene or one open reading frame, since sequence information is
CC available. This sequence represents a genomic sequence tag (GST) used in
CC the method of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 950 BP; 277 A; 184 C; 262 G; 227 T; 0 other;

Alignment Scores:

Pred. No.:	1.45e-46	Length:	950
Score:	509.00	Matches:	101
Percent Similarity:	74.86%	Conservative:	30
Best Local Similarity:	57.71%	Mismatches:	44
Query Match:	57.51%	Indels:	0
DB:	24	Gaps:	0

US-09-770-509-2 (1-178) x ABK77373 (1-950)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::||||||:::|||||      |||||||||   |||||||||      |||
Db        103 CTGCAAGGTGTTGATTTTATTGCAGTCAATACAGATGCACAAGCACTCCATTTATCAAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   ||||||      :::|||:::||||||||||||||||:::
```

```
Db        163 GCGGAAAAGAAGTTGCAACTTGGCGGCAAACTTACTCGCGGCCTTGGCGCTGGCGCCAAC 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||::::::||||||:::   :::      ::::::   :::|||
Db        223 CCAGAAATAGGCAAGAAAGCGGCAGAGGAAAGCAGAGAACAGCTGGAAGAAGTTCTTACA 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              |||:::|||:::|||||||||   ||||||||||||||||   ||||||||||||
Db        283 GGTTCTGATATGGTCTTTATTACTGCCGGAATGGGAGGAGGAACAGGCACCGGAGCAGCT 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   ||||||:::|||||||||   ||||||||||||||||||:::
Db        343 CCTGTCATTGCTGAAGTGGCAAAAGAGCTTGGCGCATTGACGGTCGGTGTTGTGACACGC 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||   :::|||:::      |||
Db        403 CCTTTCTCTTTTGAAGGCCGTAAGCGGCAAAACCAAGCGATCTCTGGGATTGCAGCGTTA 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::   ||||||||||||||||||:::||||||      ||||||   :::   |||
Db        463 AAAGAAAAAGTGGACACATTGATCGTGATTCCAAATGACCGCTTATTGGAAATGGTTGAT 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db        523 AAAAATACGCCGATGCTTGAAGCGTTCCGCGAGGCGGATAATGTATTGCGCCAAGGAGTT 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||||||||||||
Db        583 CAAGGCATTTCCGATTTAATTGCCACGCCTGGGTTGATCAACCTT 627
```

RESULT 5
AAX20250
ID AAX20250 standard; DNA; 111309 BP.
XX
AC AAX20250;
XX
DT 04-MAY-1999 (first entry)
XX
DE Borrelia burgdorferi polynucleotide sequence #3.
XX
KW Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW infection; diagnosis; characterisation; detection; ds.
XX
OS Borrelia burgdorferi.
XX
PN WO9858943-A1.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 98WO-US12764.
XX
PR 03-SEP-1997; 97US-0057483.
PR 20-JUN-1997; 97US-0050359.
PR 22-JUL-1997; 97US-0053344.
PR 22-JUL-1997; 97US-0053377.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMMUNE INC.
XX
PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
PI White OR;
XX
DR WPI; 1999-081217/07.
XX
PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of infections, particularly Lyme disease
XX
PS Claim 1; Page 738-800; 1128pp; English.
XX
CC AAX20248 to AAX20402 represent polynucleotide sequences isolated from
CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for

```
CC   the detection, diagnosis, characterisation, prevention and therapy of
CC   Bb infections, e.g. Lyme disease. They can also be used for the
CC   production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC   to a family of motile, spiral-shaped bacteria called Spirochetes.
CC   Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC   endemic relapsing fever, and Lyme borreliosis, more commonly known as
CC   Lyme disease.
XX
SQ   Sequence 111309 BP; 35956 A; 13151 C; 19075 G; 43117 T; 10 other;

Alignment Scores:
Pred. No.:              1.58e-43      Length:         111309
Score:                  506.00        Matches:        101
Percent Similarity:     74.42%        Conservative:   27
Best Local Similarity:  58.72%        Mismatches:     44
Query Match:            57.18%        Indels:         0
DB:                     20            Gaps:           0

US-09-770-509-2 (1-178) x AAX20250 (1-111309)

Qy          7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
              ||||||||||||||||||||||||||||   |||||||||      |||:::||||||
Db      12947 GTTGAATTTATTGTGGCTAATACCGATCTTCAGGCTCTCCAAACTTCTATTGCTCCCATA 13006

Qy         27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
              ||||||   ||||||      :::|||   ||||||||||||||   ||||||||||:::
Db      13007 AAAATTGCCCTTGGAGCAAAAGTTACAGCAGGGCTTGGTGCTGGGGGAAAGCCTGAGATT 13066

Qy         47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
              |||:::   :::||||||:::            |||:::      |||      ::::::
Db      13067 GGACAAGCTGCAGCAGAGGAAGACATAGATGTTATACGAAATCATCTTTCTGGTGCCGAT 13126

Qy         67 MetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValVal 86
              |||:::|||||||||   ||||||||||||||||   ||||||||||||||||||:::
Db      13127 ATGGTGTTTATTACTGCTGGTATGGGGGGCGGGACAGGAACCGGAGCAGCTCCAGTTATT 13186

Qy         87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArg 106
              |||   ||||||:::||||||||||||||||||||||||||||:::   ||||||:::
Db      13187 GCGCAAGTTGCAAAAGAGCTTGGTATTTTAACAGTTGGAGTTGTAACAAAGCCTTTTAAG 13246

Qy        107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyr 126
                 |||||||||   :::   |||||||||      |||:::      |||   |||
Db      13247 TTTGAAGGTCCTAAGAAGTTGAGACTTGCTGAGCAGGGAATAAATAACTTAAGAAAGTCT 13306

Qy        127 ValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThr 146
              ||||||||||||||:::::|||||||   ||||||   :::   ||||||   |||
Db      13307 GTAGATACATTGATCATTATTCCAAATCAAAAGCTTTTAACTGTTGTTGACAAAAGAACC 13366

Qy        147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
              |||:::   :::|||||:::   ||||||||||||||      ||||||:::|||:::
Db      13367 ACCATTAAAGATGCTTTTAAGCGTGCAGATGATGTTCTTAGAATGGGCGTTCAAGGTATT 13426

Qy        167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                    ||||||:::      |||   :::|||:::
Db      13427 GCAGGGCTTATTATTGAGCATGGAGAGGTTAATATT 13462

RESULT 6
AAX20248
ID   AAX20248 standard; DNA; 910715 BP.
XX
AC   AAX20248;
XX
DT   04-MAY-1999  (first entry)
XX
DE   Borrelia burgdorferi polynucleotide sequence #1.
XX
KW   Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW   epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW   infection; diagnosis; characterisation; detection; ds.
XX
OS   Borrelia burgdorferi.
XX
PN   WO9858943-A1.
XX
PD   30-DEC-1998.
XX
PF   18-JUN-1998;   98WO-US12764.
XX
PR   03-SEP-1997;   97US-0057483.
PR   20-JUN-1997;   97US-0050359.
PR   22-JUL-1997;   97US-0053344.
PR   22-JUL-1997;   97US-0053377.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (MEDI-) MEDIMMUNE INC.
XX
PI   Clayton R,  Dougherty BA,  Fraser C,  Lathigra R,  Smith HO;
PI   White OR;
XX
DR   WPI; 1999-081217/07.
XX
PT   New isolated Borrelia burgdorferi nucleic acids - used to develop
PT   products for the detection, diagnosis, characterisation, prevention
PT   and therapy of infections, particularly Lyme disease
XX
PS   Claim 1; Page 157-671; 1128pp; English.
XX
CC   AAX20248 to AAX20402 represent polynucleotide sequences isolated from
CC   Borrelia burgdorferi (Bb). Products derived from Bb can be used for
CC   the detection, diagnosis, characterisation, prevention and therapy of
CC   Bb infections, e.g. Lyme disease. They can also be used for the
CC   production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC   to a family of motile, spiral-shaped bacteria called Spirochetes.
CC   Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC   endemic relapsing fever, and Lyme borreliosis, more commonly known as
CC   Lyme disease.
XX
SQ   Sequence 910715 BP; 327171 A; 129646 C; 130753 G; 323091 T; 54 other;

Alignment Scores:
Pred. No.:              2.47e-42      Length:         910715
Score:                  506.00        Matches:        101
Percent Similarity:     74.42%        Conservative:   27
Best Local Similarity:  58.72%        Mismatches:     44
Query Match:            57.18%        Indels:         0
DB:                     20            Gaps:           0

US-09-770-509-2 (1-178) x AAX20248 (1-910715)

Qy          7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
              ||||||||||||||||||||||||||||   |||||||||      |||:::||||||
Db     602333 GTTGAATTTATTGTGGCTAATACCGATCTTCAGGCTCTCCAAACTTCTATTGCTCCCATA 602392

Qy         27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
              ||||||   ||||||      :::|||   ||||||||||||||   ||||||||||:::
Db     602393 AAAATTGCCCTTGGAGCAAAAGTTACAGCAGGGCTTGGTGCTGGGGGAAAGCCTGAGATT 602452

Qy         47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
              |||:::   :::||||||:::            |||:::      |||      ::::::
Db     602453 GGACAAGCTGCAGCAGAGGAAGACATAGATGTTATACGAAATCATCTTTCTGGTGCCGAT 602512

Qy         67 MetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValVal 86
              |||:::|||||||||   ||||||||||||||||   ||||||||||||||||||:::
Db     602513 ATGGTGTTTATTACTGCTGGTATGGGGGGCGGGACAGGAACCGGAGCAGCTCCAGTTATT 602572

Qy         87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArg 106
              |||   ||||||:::||||||||||||||||||||||||||||:::   ||||||:::
Db     602573 GCGCAAGTTGCAAAAGAGCTTGGTATTTTAACAGTTGGAGTTGTAACAAAGCCTTTTAAG 602632

Qy        107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyr 126
                 |||||||||   :::   |||||||||      |||:::      |||   |||
Db     602633 TTTGAAGGTCCTAAGAAGTTGAGACTTGCTGAGCAGGGAATAAATAACTTAAGAAAGTCT 602692

Qy        127 ValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThr 146
              ||||||||||||||:::::|||||||   ||||||   :::   ||||||   |||
Db     602693 GTAGATACATTGATCATTATTCCAAATCAAAAGCTTTTAACTGTTGTTGACAAAAGAACC 602752
```

```
Qy        147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
              |||:::   :::|||||||:::   |||||||||||||||      ||||||:::|||:::
Db     602753 ACCATTAAAGATGCTTTTAAGCGTGCAGATGATGTTCTTAGAATGGGCGTTCAAGGTATT 602812

Qy        167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                |||||||:::      |||   :::|||:::
Db     602813 GCAGGGCTTATTATTGAGCATGGAGAGGTTAATATT 602848

RESULT 7
ABK72889
ID   ABK72889 standard; DNA; 1134 BP.
XX
AC   ABK72889;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Bacillus licheniformis genomic sequence tag (GST) #180.
XX
KW   Differential gene expression; genomic sequenced tag; GST;
KW   altered culture condition; environmental stress;
KW   physiological provocation; ds.
XX
OS   Bacillus licheniformis.
XX
PN   WO200229113-A2.
XX
PD   11-APR-2002.
XX
PF   05-OCT-2001; 2001WO-US31437.
XX
PR   06-OCT-2000; 2000US-0680598.
PR   27-MAR-2001; 2001US-279526P.
XX
PA   (NOVO ) NOVOZYMES BIOTECH INC.
PA   (NOVO ) NOVOZYMES AS.
XX
PI   Berka R,  Clausen IG;
XX
DR   WPI; 2002-416684/44.
XX
PT   Monitoring differential expression of several genes in first Bacillus
PT   cell relative to expression of same genes in one or more second
PT   Bacillus cells, by using substrate containing Bacillus genomic
PT   sequenced tag array  -
XX
PS   Claim 4; SEQ ID NO 180; 200pp; English.
XX
CC   The invention describes a method of monitoring differential expression of
CC   genes in a first Bacillus cell relative to expression of the genes in
CC   other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC   isolated from Bacillus cells to a substrate containing array of Bacillus
CC   genomic sequenced tags (GST), examining the array, and determining
CC   relative gene expression by an observed hybridisation reporter signal of
CC   a spot in the array. The method is useful for measuring the expression of
CC   genes in a first Bacillus cell relative to expression of the same genes
CC   in one or more second Bacillus cells. The method is useful for monitoring
CC   global expression of several genes from a Bacillus cell, discovering new
CC   genes, identifying possible functions of unknown open reading frames and
CC   monitoring gene copy number variation and stability. Monitoring changes
CC   in expression of genes may be used to provide a representation of the way
CC   in which Bacillus cells adapt to changes in culture conditions,
CC   environmental stress or other physiological provocation. Extensive
CC   follow-up characterisation is unnecessary, when one spot on an array
CC   equals one gene or one open reading frame, since sequence information is
CC   available. This sequence represents a genomic sequence tag (GST) used in
CC   the method of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1134 BP; 333 A; 277 C; 307 G; 217 T; 0 other;

Alignment Scores:
Pred. No.:              1.08e-45        Length:         1134
Score:                  502.00          Matches:        97
Percent Similarity:     74.01%          Conservative:   34
Best Local Similarity:  54.80%          Mismatches:     46
Query Match:            56.72%          Indels:         0
DB:                     24              Gaps:           0

US-09-770-509-2 (1-178) x ABK72889 (1-1134)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              :::   ::::::|||||||||||||||||      |||||||||   |||||||||
Db         97 AATGACGTTCAGGGAGTCGAGTTTATCGCAGTCAACACGGATGCTCAGGCTCTCAACCTG 156

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||   |||       |||:::   :::|||      :::|||:::|||||||||||||||||
Db        157 TCAAAAGCGGAAACGAAAATGCAGATCGGTGCGAAGCTGACGCGCGGGCTCGGCGCCGGA 216

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   ||||||:::|||||||:::::::|||||:::   |||      |||:::
Db        217 GCCAATCCGGAAGTGGGCAAAAAAGCCGCAGAGGAAAGCAAAGAACAAATTGAAGAAGCG 276

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::   :::::::|||:::|||:::|||   ||||||||||||||||   ||||||
Db        277 CTGAAAGGTGCAGATATGGTGTTCGTCACAGCCGGAATGGGCGGCGGAACAGGAACGGGC 336

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||:::|||   :::|||:::::|||||   ||||||||||||||||||
Db        337 GCGGCACCTGTCATCGCACAAATCGCAAAAGATCTGGGCGCATTGACTGTCGGCGTTGTC 396

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||      |||          |||      |||:::
Db        397 ACAAGGCCGTTTACCTTTGAAGGAAGAAAAAGACAGCTTCAGGCTGCAGGCGGTATTTCA 456

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              :::   :::   ||||||||||||||||||:::||||||      ||||||   :::
Db        457 GCAATGAAGGAAGCCGTCGACACCCTGATCGTCATTCCGAATGACCGTCTTCTTGAAATC 516

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              ||||||:::|||   ||||||||||||||||||   ||||||:::|||||   :::
Db        517 GTCGATAAAAACACACCGATGCTTGAAGCGTTCCGTGAAGCGGACAACGTTCTCCGCCAA 576

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||:::::|||||||||      ||||||||||||||||||
Db        577 GGTGTTCAGGGCATTTCAGACCTGATCGCAACGCCTGGACTGATCAACCTT 627

RESULT 8
AAS52905
ID   AAS52905 standard; DNA; 1233 BP.
XX
AC   AAS52905;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Enterococcus faecalis DNA for cellular proliferation protein #333.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Enterococcus faecalis.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
```

```
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU35046.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 6542; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1233 BP; 437 A; 223 C; 270 G; 303 T; 0 other;

Alignment Scores:
Pred. No.:              7.06e-45        Length:         1233
Score:                  495.00          Matches:        100
Percent Similarity:     73.14%          Conservative:   28
Best Local Similarity:  57.14%          Mismatches:     47
Query Match:            55.93%          Indels:         0
DB:                     23              Gaps:           0

US-09-770-509-2 (1-178) x AAS52905 (1-1233)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||   |||||||||||   |||||||||      |||
Db        106 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 165

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||         |||:::|||||||||||||||||:::
Db        166 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 225

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||:::   :::      |||        ||||||
Db        226 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 285

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              :::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        286 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 345

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db        346 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 405

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||       |||:::      |||
Db        406 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCCGCTGAAGGAATTGCCTTATTA 465

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                :::   ||||||||||||:::::::::   |||       ||||||   :::   |||
Db        466 AAAGAAAACGTTGATACACTATTAATTATCTCAAACAACCGCTTATTAGAAGTCGTTGAC 525

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db        526 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 585

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::::|||||||||      ||||||   :::||||||
Db        586 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTA 630

RESULT 9
AAS51272
ID   AAS51272 standard; DNA; 1239 BP.
XX
AC   AAS51272;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Enterococcus faecalis DNA for cellular proliferation protein #49.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Enterococcus faecalis.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33413.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 3854; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
```

```
SQ   Sequence 1239 BP; 439 A; 224 C; 273 G; 303 T; 0 other;

Alignment Scores:
Pred. No.:              7.11e-45        Length:         1239
Score:                  495.00          Matches:        100
Percent Similarity:     73.14%          Conservative:   28
Best Local Similarity:  57.14%          Mismatches:     47
Query Match:            55.93%          Indels:         0
DB:                     23              Gaps:           0

US-09-770-509-2 (1-178) x AAS51272 (1-1239)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||  ||||||||||||  |||||||||      |||
Db        106 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 165

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||.        |||:::||||||||||||||||||:::
Db        166 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 225

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::||||||:::   :::      |||         ||||||
Db        226 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 285

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||||||||  ||||||||||||||||  ||||||||||
Db        286 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 345

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db        346 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 405

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||      |||:::      |||
Db        406 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCCGCTGAAGGAATTGCCTTATTA 465

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::::::::   |||      ||||||   :::   |||
Db        466 AAAGAAAACGTTGATACACTATTAATTATCTCAAACAACCGCTTATTAGAAGTCGTTGAC 525

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db        526 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 585

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::|||||||||      ||||||   :::|||||||
Db        586 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTA 630

RESULT 10
AAX13065/c
ID   AAX13065 standard; DNA; 32768 BP.
XX
AC   AAX13065;
XX
DT   19-MAR-1999  (first entry)
XX
DE   Enterococcus faecalis genome contig SEQ ID NO:128.
XX
KW   Enterococcus faecalis; contig; detection; Enterococcal infection;
KW   vaccine; attenuation; computer readable medium; ds.
XX
OS   Enterococcus faecalis.
XX
PN   WO9850555-A2.
XX
PD   12-NOV-1998.
XX
PF   04-MAY-1998;   98WO-US08985.
XX
PR   14-NOV-1997;   97US-0066009.
PR   06-MAY-1997;   97US-0044031.
PR   16-MAY-1997;   97US-0046655.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Dillon PJ,  Kunsch CA;
XX
DR   WPI; 1999-045171/04.
XX
PT   New isolated Enterococcus faecalis polynucleotides and polypeptides
PT   - used to develop products for the detection of Enterococcus and for
PT   use in vaccines for prevention or attenuation of Enterococcus
PT   infection.
XX
PS   Claim 1; Page 767-783; 2084pp; English.
XX
CC   A computer readable medium has been developed which has recorded on it
CC   982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC   AAX12938 to AAX13919 represent these nucleotide sequences which are
CC   primary nucleotide sequences, also known as contigs. The computer-based
CC   system can identify fragments of the Enterococcus faecalis genome with
CC   commercial importance. The products can be used to detect the presence
CC   of Enterococcus faecalis in samples. They can also be used for
CC   diagnosing Enterococcal infection in an animal and monitoring
CC   progression of disease, and for identifying agents which can be used to
CC   modulate the growth or pathogenicity of Enterococcus faecalis, or
CC   another related organism, in vivo or in vitro. In particular the
CC   polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC   can be used in vaccines to prevent or attenuate an Enterococcal
CC   infection.
XX
SQ   Sequence 32768 BP; 9529 A; 6883 C; 5611 G; 10720 T; 25 other;

Alignment Scores:
Pred. No.:              5.17e-43        Length:         32768
Score:                  495.00          Matches:        100
Percent Similarity:     73.14%          Conservative:   28
Best Local Similarity:  57.14%          Mismatches:     47
Query Match:            55.93%          Indels:         0
DB:                     20              Gaps:           0

US-09-770-509-2 (1-178) x AAX13065 (1-32768)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||  ||||||||||||  |||||||||      |||
Db      27971 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 27912

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||.        |||:::||||||||||||||||||:::
Db      27911 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 27852

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::||||||:::   :::      |||         ||||||
Db      27851 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 27792

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||||||||  ||||||||||||||||  ||||||||||
Db      27791 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 27732

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db      27731 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 27672

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   ||||||         |||:::      |||
Db      27671 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTGCNTGCTGAAGGAATTGCCTTATTA 27612

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::::::::   |||      ||||||   :::   |||
Db      27611 AAAGAAAACGTTGATACACTATTAATTATCTCAAATAACCGCTTATTAGAAGTCGTTGAC 27552

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db      27551 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 27492

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
```

```
                ::::|||::::::||||||||||      ||||||   :::|||||||
Db      27491 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTG 27447

RESULT 11
ABA90521/c
ID   ABA90521 standard; DNA; 2365589 BP.
XX
AC   ABA90521;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Genomic sequence of Lactococcus lactis IL1403.
XX
KW   Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.
XX
OS   Lactococcus lactis IL1403.
XX
PN   FR2807446-A1.
XX
PD   12-OCT-2001.
XX
PF   11-APR-2000; 2000FR-0004630.
XX
PR   11-APR-2000; 2000FR-0004630.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI   Bolotine A,  Sorokine A,  Renault P,  Ehrlich SD;
XX
DR   WPI; 2002-043418/06.
XX
PT   New nucleotide sequence useful in the identification or Lactococcus
PT   lactis and related species -
XX
PS   Claim 1; SEQ ID 1; 2504pp; French.
XX
CC   The present invention is related to a Lactococcus lactis nucleotide
CC   sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC   nucleic acid sequence is useful in the detection and/or amplification of
CC   nucleic acid sequence, particularly to identify Lactococcus lactis or
CC   related species. The proteins of the invention are useful for the
CC   biosynthesis or biodegradation of a composition of interest. The
CC   invention helps research in lactic bacteria, particularly useful in the
CC   production of yogurt and cheese.
CC   Note: The sequence data for this patent is based on equivalent patent
CC   WO200177334 (published 18-OCT-2001) which is available in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2365589 BP; 765914 A; 415261 C; 420487 G; 763927 T; 0 other;

Alignment Scores:
Pred. No.:              2.32e-40        Length:         2365589
Score:                  493.00          Matches:        98
Percent Similarity:     72.57%          Conservative:   29
Best Local Similarity:  56.00%          Mismatches:     48
Query Match:            55.71%          Indels:         0
DB:                     24              Gaps:           0

US-09-770-509-2 (1-178) x ABA90521 (1-2365589)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              :::   |||||||||||||||   ||||||||||   |||||||||      |||
Db    1938822 GTTTCTGGCGTTGAATTTATCGCCGCTAATACTGATGTACAAGCATTGCGTAGCTCAAAA 1938763

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||        |||   ||||||      :::|||:::||||||||||||||||:::::
Db    1938762 GCAGATACAGTTATTCAACTTGGGCCAAAATTGACTCGCGGTTTGGGTGCTGGTGCGCAA 1938703

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||||||||:::|||||:::           :::   :::   |||:::
Db    1938702 CCTGAAGTTGGTAAGCGCGCAGCTGAAGAATCAGCTGAAACAGTATCACAAGCCCTTGAA 1938643

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||:::|||:::|||||||||   ||||||||||||||||   ||||||||||||
Db    1938642 GGCTCAGATATGATTTTCATCACTGCCGGTATGGGTGGTGGAACTGGTACAGGTGCTGCA 1938583

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   :::|||:::|||||||||   |||||||||||||||||:::
Db    1938582 CCAGTTATCGCACAAATTGCTAAAGAGCTTGGCGCTTTGACTGTGGGTGTTGTTACTCGT 1938523

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||:::      |||      |||::::::   |||
Db    1938522 CCATTTGGTTTTGAAGGATCAAAACGTTCATACTTTGCAACTGAAGGAATTGAAGCTCTT 1938463

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                    ||||||||||||:::::::::   |||   |||||||||   :::   |||
Db    1938462 CGTGCAAATGTTGATACACTTCTCATTATTTCAAACAATAACTTGCTTGAAATTGTTGAT 1938403

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   :::   ||||||   |||   ||||||:::||||||   :::||||||
Db    1938402 AAGAAAACTCCTTTAACTGAAGCTTTGCGTGAAGCTGATAATGTTTTACGTCAAGGTGTT 1938343

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::||||||||||||||||||      ||||||:::|||||||||
Db    1938342 CAAGGGGTAACTGACTTGATTACTAACCCAGGAATGATTAACCTT 1938298

RESULT 12
AAH52021
ID   AAH52021 standard; DNA; 1140 BP.
XX
AC   AAH52021;
XX
DT   04-SEP-2001  (first entry)
XX
DE   Mycobacterium tuberculosis potential drug target gene SEQ ID 75.
XX
KW   Drug target; growth; organism viability; characterisation; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO200135317-A1.
XX
PD   17-MAY-2001.
XX
PF   13-NOV-2000; 2000WO-US31152.
XX
PR   12-NOV-1999;   99US-0165086.
PR   12-NOV-1999;   99US-0165124.
PR   01-FEB-2000; 2000US-0179531.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Eisenberg D,  Rotstein SH,  Marcotte EM;
XX
DR   WPI; 2001-329193/34.
DR   P-PSDB; AAG81170.
XX
PT   Identifying nucleotide or polypeptide sequence for use as drug target,
PT   involves providing algorithm that analyzes a functional relationship
PT   between nucleotide or polypeptide sequences, and comparing the
PT   sequences  -
XX
PS   Disclosure; Page 97; 207pp; English.
XX
CC   This invention relates to a method for identifying a nucleotide or
CC   polypeptide sequence that may be a drug target, or essential for growth
CC   or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
CC   represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium
CC   tuberculosis proteins which are potential drug targets. The DNA and
CC   protein sequences are used to illustrate the method of the invention. The
CC   method involves providing an unknown nucleotide or polypeptide sequences,
CC   and comparing it to a number of sequences along with at least one
CC   algorithm capable of analysing a functional relationship between
CC   nucleotide and polypeptide sequences. The method is useful for
CC   characterising the function of nucleic acids and polypeptides that may be
CC   useful as a target for a drug or essential for the growth or viability of
```

```
CC   an organism.
XX
SQ   Sequence 1140 BP; 203 A; 368 C; 388 G; 181 T; 0 other;

Alignment Scores:
Pred. No.:              8.02e-44        Length:         1140
Score:                  485.00          Matches:        95
Percent Similarity:     74.29%          Conservative:   35
Best Local Similarity:  54.29%          Mismatches:     45
Query Match:            54.80%          Indels:         0
DB:                     22              Gaps:           0

US-09-770-509-2 (1-178) x AAH52021 (1-1140)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||||||||||||      |||||||||   |||||||||      |||
Db         94 CTCAAAGGCGTGGAATTCATCGCGATCAACACCGACGCCCAGGCGTTGTTGATGAGCGAT 153

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||:::|||   |||:::||||||||||||||:::
Db        154 GCCGACGTCAAACTCGACGTCGGCCGCGACTCCACCCGCGGGCTGGGCGCCGGCGCCGAT 213

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||      |||   :::|||:::   :::|||:::
Db        214 CCGGAGGTCGGCCGTAAGGCCGCCGAGGACGCCAAGGACGAGATCGAAGAGCTGCTGCGC 273

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||:::|||   |||   ||||||||||||   ||||||   |||
Db        274 GGTGCCGACATGGTGTTTGTCACCGCCGGCGAGGGGGGCGGAACCGGCACCGGGGGGGCA 333

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||||:::||||||:::|||||||   ||||||||||||||||||:::
Db        334 CCCGTCGTCGCCAGCATCGCCCGCAAGCTGGGCGCGTTGACCGTCGGTGTGGTCACCCGG 393

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||:::      |||      |||:::      |||
Db        394 CCGTTCTCGTTCGAGGGCAAGCGACGCAGCAATCAGGCCGAAAATGGCATCGCGGCGCTG 453

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::      ||||||||||||||||:::||||||      ||||||   :::   |||
Db        454 CGGGAGAGTTGCGACACCCTCATCGTGATTCCCAACGACCGGTTGCTGCAGATGGGAGAT 513

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
                 :::   ::::::::::::|||||||||   ||||||:::|||||||||   ||||||
Db        514 GCCGCGGTATCGCTGATGGATGCTTTCCGTAGCGCCGACGAGGTGCTGCTCAACGGCGTG 573

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::||||||||||||      ||||||||||||||:::
Db        574 CAGGGCATCACCGACCTGATTACCACCCCGGGTCTAATCAACGTC 618

RESULT 13
AAI99683/c
ID   AAI99683 standard; DNA; 4403765 BP.
XX
AC   AAI99683;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 4; SEQ ID NO 2; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4403765 BP; 757105 A; 1447799 C; 1441301 G; 757371 T; 189 other;

Alignment Scores:
Pred. No.:              3.97e-39        Length:         4403765
Score:                  485.00          Matches:        95
Percent Similarity:     74.29%          Conservative:   35
Best Local Similarity:  54.29%          Mismatches:     45
Query Match:            54.80%          Indels:         0
DB:                     22              Gaps:           0

US-09-770-509-2 (1-178) x AAI99683 (1-4403765)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||||||||||||      |||||||||   |||||||||      |||
Db    2408087 CTCAAAGGCGTGGAATTCATCGCGATCAACACCGACGCCCAGGCGTTGTTGATGAGCGAT 2408028

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||:::|||   |||:::||||||||||||||:::
Db    2408027 GCCGACGTCAAACTCGACGTCGGCCGCGACTCCACCCGCGGGCTGGGCGCCGGCGCCGAT 2407968

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||      |||   :::|||:::   :::|||:::
Db    2407967 CCGGAGGTCGGCCGTAAGGCCGCCGAGGACGCCAAGGACGAGATCGAAGAGCTGCTGCGC 2407908

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||:::|||   |||   ||||||||||||   ||||||   |||
Db    2407907 GGTGCCGACATGGTGTTTGTCACCGCCGGCGAGGGGGGCGGAACCGGCACCGGGGGGGCA 2407848

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||||:::||||||:::|||||||   ||||||||||||||||||:::
Db    2407847 CCCGTCGTCGCCAGCATCGCCCGCAAGCTGGGCGCGTTGACCGTCGGTGTGGTCACCCGG 2407788

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||:::      |||      |||:::      |||
Db    2407787 CCGTTCTCGTTCGAGGGCAAGCGACGCAGCAATCAGGCCGAAAATGGCATCGCGGCGCTG 2407728

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::      ||||||||||||||||:::||||||      ||||||   :::   |||
Db    2407727 CGGGAGAGTTGCGACACCCTCATCGTGATTCCCAACGACCGGTTGCTGCAGATGGGAGAT 2407668

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
                 :::   ::::::::::::|||||||||   ||||||:::|||||||||   ||||||
Db    2407667 GCCGCGGTATCGCTGATGGATGCTTTCCGTAGCGCCGACGAGGTGCTGCTCAACGGCGTG 2407608

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::||||||||||||      ||||||||||||||:::
```

```
Db    2407607 CAGGGCATCACCGACCTGATTACCACCCCGGGTCTAATCAACGTC 2407563

RESULT 14
AAI99682/c
ID   AAI99682 standard; DNA; 4411529 BP.
XX
AC   AAI99682;
XX
DT   15-JAN-2002  (first entry)
XX
DE   Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX
KW   Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW   variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX
OS   Mycobacterium tuberculosis.
XX
PN   US6294328-B1.
XX
PD   25-SEP-2001.
XX
PF   24-JUN-1998;   98US-0103840.
XX
PR   24-JUN-1998;   98US-0103840.
XX
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fleischmann RD,  White OR,  Fraser CM,  Venter JC;
XX
DR   WPI; 2001-647261/74.
XX
PT   Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT   determining the nucleotide sequence of the strain at positions in the
PT   genome corresponding to positions where M. tuberculosis strains CDC
PT   1551 and H37Rv differ  -
XX
PS   Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
XX
CC   The invention relates to evaluating strain variation within and between
CC   different populations of the tuberculosis bacterial pathogen,
CC   Mycobacterium tuberculosis or related Mycobacterium by determining the
CC   nucleotide sequence of the first strain at positions in the complete
CC   sequence of the genome that correspond to positions that differ in the
CC   nucleotide sequences of M. tuberculosis strains CDC 1551 (AAI99683) and
CC   H37Rv (AAI99682). The method is useful for evaluating strain variation of
CC   M. tuberculosis and has valuable application in the fields of
CC   tuberculosis genetics, epidemiology, patient treatment and epidemic
CC   monitoring.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=6294328B1.
XX
SQ   Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;

Alignment Scores:
Pred. No.:              3.98e-39      Length:          4411529
Score:                  485.00        Matches:         95
Percent Similarity:     74.29%        Conservative:    35
Best Local Similarity:  54.29%        Mismatches:      45
Query Match:            54.80%        Indels:          0
DB:                     22            Gaps:            0

US-09-770-509-2 (1-178) x AAI99682 (1-4411529)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||||||||||||      |||||||||   |||||||||      |||
Db    2409429 CTCAAAGGCGTGGAATTCATCGCGATCAACACCGACGCCCAGGCGTTGTTGATGAGCGAT 2409370

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||:::|||   |||:::||||||||||||||:::
Db    2409369 GCCGACGTCAAACTCGACGTCGGCCGCGACTCCACCCGCGGGCTGGGCGCCGGCGCCGAT 2409310

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||      |||   :::|||:::   :::|||:::
Db    2409309 CCGGAGGTCGGCCGTAAGGCCGCCGAGGACGCCAAGGACGAGATCGAAGAGCTGCTGCGC 2409250

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||:::||| |||  ||||||||||||  ||||||  |||
Db    2409249 GGTGCCGACATGGTGTTTGTCACCGCCGGCGAGGGGGGCGGAACCGGCACCGGGGGGGCA 2409190

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              |||||||||||||||:::||||||:::||||||  ||||||||||||||||||:::
Db    2409189 CCCGTCGTCGCCAGCATCGCCCGCAAGCTGGGCGCGTTGACCGTCGGTGTGGTCACCCGG 2409130

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||:::      |||      |||:::      |||
Db    2409129 CCGTTCTCGTTCGAGGGCAAGCGACGCAGCAATCAGGCCGAAAATGGCATCGCGGCGCTG 2409070

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::      ||||||||||||||||:::|||||||      ||||||  :::  |||
Db    2409069 CGGGAGAGTTGCGACACCCTCATCGTGATTCCCAACGACCGGTTGCTGCAGATGGGAGAT 2409010

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              :::   ::::::::::::|||||||||  ||||||:::|||||||||  ||||||
Db    2409009 GCCGCGGTATCGCTGATGGATGCTTTCCGTAGCGCCGACGAGGTGCTGCTCAACGGCGTG 2408950

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::||||||||||||      |||||||||||||||:::
Db    2408949 CAGGGCATCACCGACCTGATTACCACCCCGGGTCTAATCAACGTC 2408905

RESULT 15
AAS52133
ID   AAS52133 standard; DNA; 1172 BP.
XX
AC   AAS52133;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus DNA for cellular proliferation protein #550.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU34274.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 4715; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
```

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1172 BP; 390 A; 194 C; 260 G; 328 T; 0 other;

Alignment Scores:

Pred. No.:	2.95e-43	Length:	1172
Score:	480.00	Matches:	95
Percent Similarity:	73.14%	Conservative:	33
Best Local Similarity:	54.29%	Mismatches:	47
Query Match:	54.24%	Indels:	0
DB:	23	Gaps:	0

US-09-770-509-2 (1-178) x AAS52133 (1-1172)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              :::      ||||||||||||      |||||||||   |||||||||      |||
Db        103 ATGAATAATGTTGAATTTATCGCTATCAACACAGACGGTCAAGCTTTAAACTTATCTAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      ||||||   :::|||:::   :::|||:::||||||||||||||||:::
Db        163 GCTGAATCTAAAATCCAAATCGGTGAAAAATTAACACGTGGTTTAGGAGCAGGAGCTAAT 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||:::::|||||||:::   :::      |||:::      :::|||
Db        223 CCTGAAATCGGTAAAAAAGCTGCAGAGGAATCTCGTGAACAAATTGAAGATGCAATCCAA 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||:::|||   ||||||||||||||||||   ||||||||||||
Db        283 GGTGCAGACATGGTATTTGTTACTTCTGGTATGGGTGGCGGAACTGGTACTGGTGCAGCA 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||:::|||   ||||||||||||||||||:::
Db        343 CCAGTCGTTGCTAAAATTGCAAAAGAAATGGGCGCATTAACTGTTGGTGTTGTAACTCGT 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||      |||   |||||||||:::   :::
Db        403 CCATTTAGTTTTGAAGGACGTAAACGTCAAACTCAAGCTGCTGCTGGAGTAGAAGCTATG 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                  ||||||||||||||||||:::||||||      ||||||   :::   |||
Db        463 AAAGCTGCAGTAGATACATTAATCGTTATACCAAATGACCGTTTATTAGATATCGTTGAC 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||||||||   |||:::||||||||:::   ||||||:::||||||   :::||||||
Db        523 AAATCTACGCCAATGATGGAAGCATTTAAAGAAGCTGACAACGTGTTACGCCAAGGTGTA 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::::|||||||||         |||   :::||||||
Db        583 CAAGGTATCTCAGACTTAATCGCTGTTTCTGGTGAAGTAAACTTA 627
```

Search completed: June 2, 2003, 14:34:50
Job time : 20511 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:48:49 ; Search time 57 Seconds
(without alignments)
957.692 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09770509/runat_27052003_083650_18840/app_query.fasta_1.327
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09770509_@CGN_1_1_40_@runat_27052003_083650_18840 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	485	54.8	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	2	485	54.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	3	480	54.2	1185	4	US-09-134-001C-1997	Sequence 1997, Ap
	4	472	53.3	1005	4	US-08-987-146-3	Sequence 3, Appli
	5	472	53.3	1008	4	US-08-987-146-1	Sequence 1, Appli
	6	472	53.3	1260	4	US-09-120-426-1	Sequence 1, Appli
	7	472	53.3	2702	4	US-08-987-146-4	Sequence 4, Appli
	8	472	53.3	21706	4	US-08-961-527-36	Sequence 36, Appl
	9	464	52.4	7143	4	US-09-381-862-4	Sequence 4, Appli
	10	463	52.3	1628	2	US-08-883-515-3	Sequence 3, Appli
	11	429	48.5	1425	2	US-08-883-515-1	GENERAL INFORMAT
c	12	376	42.5	396	4	US-09-134-001C-1996	Sequence 1996, Ap
	13	369.5	41.8	665	4	US-09-221-017B-658	Sequence 658, App
	14	271	30.6	669	4	US-09-120-426-3	Sequence 3, Appli
	15	196	22.1	850	3	US-08-961-083-123	Sequence 123, App
	16	90.5	10.2	7100	4	US-09-308-375-1	Sequence 1, Appli
c	17	85.5	9.7	2245	2	US-08-617-801A-3	Sequence 3, Appli
	18	83	9.4	1888	1	US-08-229-145-13	Sequence 13, Appl
	19	81.5	9.2	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	20	81.5	9.2	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	21	80.5	9.1	1621	3	US-09-013-881-14	Sequence 14, Appl
	22	79	8.9	1167	4	US-09-134-001C-2114	Sequence 2114, Ap
	23	78.5	8.9	1574	4	US-09-189-527-10	Sequence 10, Appl
	24	78.5	8.9	9244	4	US-08-961-527-68	Sequence 68, Appl
c	25	78	8.8	2109	2	US-08-617-801A-5	Sequence 5, Appli
	26	78	8.8	8906	2	US-08-826-267-1	Sequence 1, Appli
	27	77.5	8.8	1946	2	US-08-861-464-1	Sequence 1, Appli
	28	77.5	8.8	1946	2	US-08-396-001-1	Sequence 1, Appli
	29	77.5	8.8	1946	4	US-09-323-433A-1	Sequence 1, Appli
	30	76.5	8.6	1269	4	US-09-134-001C-197	Sequence 197, App
c	31	76.5	8.6	2844	4	US-09-257-703-2	Sequence 2, Appli
c	32	76.5	8.6	3156	2	US-08-887-518-1	Sequence 1, Appli
c	33	76.5	8.6	3156	2	US-09-023-321-1	Sequence 1, Appli
c	34	76.5	8.6	3156	2	US-09-032-475-1	Sequence 1, Appli
c	35	75.5	8.5	1229	1	US-08-440-861-1	Sequence 1, Appli
c	36	75.5	8.5	1229	1	US-08-433-854-1	Sequence 1, Appli
c	37	75.5	8.5	1229	1	US-08-174-745A-1	Sequence 1, Appli
c	38	75.5	8.5	1229	2	US-08-195-947-1	Sequence 1, Appli
c	39	75.5	8.5	1229	2	US-08-433-885-1	Sequence 1, Appli
c	40	75.5	8.5	1229	2	US-08-433-908B-1	Sequence 1, Appli
c	41	75.5	8.5	1229	4	US-08-410-614-1	Sequence 1, Appli
c	42	75	8.5	2504	1	US-08-073-384C-3	Sequence 3, Appli
c	43	75	8.5	2504	1	US-08-254-359A-3	Sequence 3, Appli
c	44	75	8.5	2504	1	US-08-483-043-3	Sequence 3, Appli
c	45	75	8.5	2504	1	US-08-481-238-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:

Pred. No.:	1.02e-46	Length:	4403765
Score:	485.00	Matches:	95
Percent Similarity:	74.29%	Conservative:	35
Best Local Similarity:	54.29%	Mismatches:	45
Query Match:	54.80%	Indels:	0
DB:	4	Gaps:	0

US-09-770-509-2 (1-178) x US-09-103-840A-2 (1-4403765)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||||||||||||      |||||||||  |||||||||      |||
Db    2408087 CTCAAAGGCGTGGAATTCATCGCGATCAACACCGACGCCCAGGCGTTGTTGATGAGCGAT 2408028

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::  :::|||:::|||  |||:::||||||||||||||||:::
Db    2408027 GCCGACGTCAAACTCGACGTCGGCCGCGACTCCACCCGCGGGCTGGGCGCCGGCGCCGAT 2407968

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||      |||   :::|||:::   :::|||:::
Db    2407967 CCGGAGGTCGGCCGTAAGGCCGCCGAGGACGCCAAGGACGAGATCGAAGAGCTGCTGCGC 2407908

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              :::::|||:::|||:::|||   |||   ||||||||||||   ||||||   |||
Db    2407907 GGTGCCGACATGGTGTTTGTCACCGCCGGCGAGGGGGGCGGAACCGGCACCGGGGGGGCA 2407848

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||||:::|||||:::||||||   ||||||||||||||||||||:::
Db    2407847 CCCGTCGTCGCCAGCATCGCCCGCAAGCTGGGCGCGTTGACCGTCGGTGTGGTCACCCGG 2407788

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||:::      |||      |||:::      |||
Db    2407787 CCGTTCTCGTTCGAGGGCAAGCGACGCAGCAATCAGGCCGAAAATGGCATCGCGGCGCTG 2407728

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::      ||||||||||||||:::||||||      ||||||   :::   |||
Db    2407727 CGGGAGAGTTGCGACACCCTCATCGTGATTCCCAACGACCGGTTGCTGCAGATGGGAGAT 2407668

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              :::   ||||||||||||||||||||   ||||||:::||||||||   ||||||
Db    2407667 GCCGCGGTATCGCTGATGGATGCTTTCCGTAGCGCCGACGAGGTGCTGCTCAACGGCGTG 2407608

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::||||||||||||      ||||||||||||||:::
Db    2407607 CAGGGCATCACCGACCTGATTACCACCCCGGGTCTAATCAACGTC 2407563

RESULT 2
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:              1.02e-46      Length:         4411529
Score:                  485.00        Matches:        95
Percent Similarity:     74.29%        Conservative:   35
Best Local Similarity:  54.29%        Mismatches:     45
Query Match:            54.80%        Indels:         0
DB:                     4             Gaps:           0

US-09-770-509-2 (1-178) x US-09-103-840A-1 (1-4411529)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||||||||||||      |||||||||  |||||||||      |||
Db    2409429 CTCAAAGGCGTGGAATTCATCGCGATCAACACCGACGCCCAGGCGTTGTTGATGAGCGAT 2409370

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::  :::|||:::|||  |||:::||||||||||||||||:::
Db    2409369 GCCGACGTCAAACTCGACGTCGGCCGCGACTCCACCCGCGGGCTGGGCGCCGGCGCCGAT 2409310

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||      |||   :::|||:::   :::|||:::
Db    2409309 CCGGAGGTCGGCCGTAAGGCCGCCGAGGACGCCAAGGACGAGATCGAAGAGCTGCTGCGC 2409250

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              :::::|||:::|||:::|||   |||   ||||||||||||   ||||||   |||
Db    2409249 GGTGCCGACATGGTGTTTGTCACCGCCGGCGAGGGGGGCGGAACCGGCACCGGGGGGGCA 2409190

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||||:::|||||:::||||||   ||||||||||||||||||||:::
Db    2409189 CCCGTCGTCGCCAGCATCGCCCGCAAGCTGGGCGCGTTGACCGTCGGTGTGGTCACCCGG 2409130

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||:::      |||      |||:::      |||
Db    2409129 CCGTTCTCGTTCGAGGGCAAGCGACGCAGCAATCAGGCCGAAAATGGCATCGCGGCGCTG 2409070

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::      ||||||||||||||:::||||||      ||||||   :::   |||
Db    2409069 CGGGAGAGTTGCGACACCCTCATCGTGATTCCCAACGACCGGTTGCTGCAGATGGGAGAT 2409010

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              :::   ||||||||||||||||||||   ||||||:::||||||||   ||||||
Db    2409009 GCCGCGGTATCGCTGATGGATGCTTTCCGTAGCGCCGACGAGGTGCTGCTCAACGGCGTG 2408950

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::||||||||||||      ||||||||||||||:::
Db    2408949 CAGGGCATCACCGACCTGATTACCACCCCGGGTCTAATCAACGTC 2408905

RESULT 3
US-09-134-001C-1997
; Sequence 1997, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1997
;   LENGTH: 1185
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1997

Alignment Scores:
Pred. No.:              2.32e-51      Length:         1185
Score:                  480.00        Matches:        95
Percent Similarity:     73.14%        Conservative:   33
Best Local Similarity:  54.29%        Mismatches:     47
Query Match:            54.24%        Indels:         0
DB:                     4             Gaps:           0

US-09-770-509-2 (1-178) x US-09-134-001C-1997 (1-1185)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              :::      ||||||||||||      |||||||||  |||||||||      |||
Db        103 ATGAATAATGTAGAATTTATTGCCATCAACACTGATGGACAAGCTTTAAACTTATCAAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      ||||||  :::|||:::  :::|||:::||||||||||||||||:::
Db        163 GCTGAATCAAAAATTCAAATTGGTGAGAAATTAACACGTGGATTAGGTGCTGGTGCTAAT 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
```

```
                ||||||:::||||||::::::||||||:::   :::      |||:::      :::|||
Db        223 CCTGAAATTGGGAAGAAAGCAGCAGAAGAATCAAGAGAACAAATTGAAGACGCTATCCAA 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
               ::::::|||:::|||:::|||   ||||||||||||||||   ||||||||||||
Db        283 GGTGCTGATATGGTATTCGTAACTGCTGGTATGGGTGGCGGTACTGGAACAGGTGCTGCA 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
               ||||||||||||   :::|||:::|||:::|||   ||||||||||||||||||:::
Db        343 CCGGTTGTTGCTAAAATAGCAAAAGAAATGGGTGCTTTAACTGTAGGTGTTGTTACGCGT 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
               ||||||      ||||||      |||      |||   ||||||||||:::   :::
Db        403 CCATTCGGTTTCGAAGGTCGTAAGCGCCAAACACAAGCGGCAGCTGGCGTAGAATCTATG 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
               ||||||||||||||||||:::||||||      ||||||   :::   |||
Db        463 AAAGCAGCAGTGGATACATTAATTGTTATTCCAAATGATCGCTTATTAGATATCGTTGAC 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
               |||||||||   |||:::||||||||||:::   ||||||:::||||||   :::||||||
Db        523 AAATCTACGCCAATGATGGAAGCATTTAAAGAAGCGGATAATGTATTACGTCAAGGTGTA 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
               :::|||:::::::|||||||||          |||   :::||||||
Db        583 CAAGGTATTTCAGATTTAATTGCAGTATCAGGTGAAGTGAACTTA 627

RESULT 4
US-08-987-146-3
; Sequence 3, Application US/08987146
; Patent No. 6350866
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Rockey, Pamela K.
;    APPLICANT:  Wang, Q. May
;    APPLICANT:  Rosteck Jr., Paul R.
;    TITLE OF INVENTION:  Streptococcus Pneumoniae Gene Sequence
;    TITLE OF INVENTION:  FtsZ
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/987,146
;      FILING DATE:  December 8, 1997
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11755
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1005 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-987-146-3
```

```
Alignment Scores:
Pred. No.:                1.87e-50        Length:          1005
Score:                    472.00          Matches:         91
Percent Similarity:       72.83%          Conservative:    35
Best Local Similarity:    52.60%          Mismatches:      47
Query Match:              53.33%          Indels:          0
DB:                       4               Gaps:            0

US-09-770-509-2 (1-178) x US-08-987-146-3 (1-1005)

Qy          6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
              ||||||||||||||||   ||||||||||||   |||||||||      :::   |||
Db        112 GGCGUAGAAUUUAUCGCAGCAAACACAGAUGUACAAGCAUUGAGUAGUACAAAAGCUGAG 171

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
                  |||   ||||||     :::|||:::||||||||||||||||   :::|||||||
Db        172 ACUGUUAUUCAGUUGGGACCUAAAUUGACUCGUGGUUUGGGUGCAGGAGGUCAACCUGAG 231

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
              :::|||:::::::::||||||:::   :::      :::         :::      :::
Db        232 GUUGGUCGUAAAGCCGCUGAAGAAAGCGAAGAAACACUGACGGAAGCUAUUAGUGGUGCC 291

Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
              :::|||:::||||||||   ||||||||||||||:::   ||||||||||||||||||
Db        292 GAUAUGGUCUUCAUCACUGCUGGUAUGGGAGGAGGCUCUGGAACUGGAGCUGCUCCUGUU 351

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
              :::|||   :::|||:::::::||||||   ||||||||||||||||||:::   ||||||
Db        352 AUUGCUCGUAUCGCCAAAGAUUUAGGUGCGCUUACAGUUGGUGUUGUAACACGUCCCUUU 411

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
                    ||||||      |||   :::   |||      |||:::   :::|||   :::
Db        412 GGUUUUGAAGGAAGUAAGCGUGGACAAUUUGCUGUAGAAGGAAUCAAUCAACUUCGUGAG 471

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
              :::||||||||||||||:::::::::   |||   |||||||||   :::   ||||||
Db        472 CAUGUAGACACUCUAUUGAUUAUCUCAAACAACAAUUUGCUUGAAAUUGUUGAUAAGAAA 531

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
              |||   :::|||||||||          ||||||:::||||||   :::||||||:::|||
Db        532 ACACCGCUUUUGGAGGCUCUUAGCGAAGCGGAUAACGUUCUUCGUCAAGGUGUUCAAGGG 591

Qy        166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::||||||||||||      ||||||||||||||||||
Db        592 AUUACCGAUUUGAUUACCAAUCCAGGAUUGAUUAACCUU 630

RESULT 5
US-08-987-146-1
; Sequence 1, Application US/08987146
; Patent No. 6350866
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Rockey, Pamela K.
;    APPLICANT:  Wang, Q. May
;    APPLICANT:  Rosteck Jr., Paul R.
;    TITLE OF INVENTION:  Streptococcus Pneumoniae Gene Sequence
;    TITLE OF INVENTION:  FtsZ
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/987,146
```

```
;      FILING DATE:  December 8, 1997
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11755
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1008 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1005
US-08-987-146-1

Alignment Scores:
Pred. No.:                1.88e-50      Length:          1008
Score:                    472.00        Matches:         91
Percent Similarity:       72.83%        Conservative:    35
Best Local Similarity:    52.60%        Mismatches:      47
Query Match:              53.33%        Indels:          0
DB:                       4             Gaps:            0

US-09-770-509-2 (1-178) x US-08-987-146-1 (1-1008)

Qy          6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
              ||||||||||||||||   ||||||||||||   |||||||||      :::   |||
Db        112 GGCGTAGAATTTATCGCAGCAAACACAGATGTACAAGCATTGAGTAGTACAAAAGCTGAG 171

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
                 |||   ||||||     :::|||:::||||||||||||||||   :::||||||
Db        172 ACTGTTATTCAGTTGGGACCTAAATTGACTCGTGGTTTGGGTGCAGGAGGTCAACCTGAG 231

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
              :::|||::::::::::||||||:::   :::      :::         :::      :::
Db        232 GTTGGTCGTAAAGCCGCTGAAGAAAGCGAAGAAACACTGACGGAAGCTATTAGTGGTGCC 291

Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
              :::|||:::||||||||   ||||||||||||:::   ||||||||||||||||||
Db        292 GATATGGTCTTCATCACTGCTGGTATGGGAGGAGGCTCTGGAACTGGAGCTGCTCCTGTT 351

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
              :::|||   :::|||:::::|||||   ||||||||||||||||:::   ||||||
Db        352 ATTGCTCGTATCGCCAAAGATTTAGGTGCGCTTACAGTTGGTGTTGTAACACGTCCCTTT 411

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
                 ||||||      |||   :::   |||      |||:::   :::|||   :::
Db        412 GGTTTTGAAGGAAGTAAGCGTGGACAATTTGCTGTAGAAGGAATCAATCAACTTCGTGAG 471

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
              :::||||||||||||:::::::::   |||   |||||||||   :::   ||||||
Db        472 CATGTAGACACTCTATTGATTATCTCAAACAACAATTTGCTTGAAATTGTTGATAAGAAA 531

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
              |||   :::||||||||||         ||||||:::||||||   :::|||||||:::|||
Db        532 ACACCGCTTTTGGAGGCTCTTAGCGAAGCGGATAACGTTCTTCGTCAAGGTGTTCAAGGG 591

Qy        166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::||||||||||||      ||||||||||||||||||
Db        592 ATTACCGATTTGATTACCAATCCAGGATTGATTAACCTT 630

RESULT 6
US-09-120-426-1
; Sequence 1, Application US/09120426
; Patent No. 6197300
; GENERAL INFORMATION:
;  APPLICANT: Fueyo, Joanna Lynn
;  APPLICANT: Lonetto, Michael A.
;  APPLICANT: Pearce, Kenneth
;  TITLE OF INVENTION: ftsZ
;  FILE REFERENCE: GM10068
;  CURRENT APPLICATION NUMBER: US/09/120,426
;  CURRENT FILING DATE: 1998-07-22
;  EARLIER APPLICATION NUMBER: 60/055,720
;  EARLIER FILING DATE: 1997-08-12
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 1260
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
US-09-120-426-1

Alignment Scores:
Pred. No.:                2.61e-50      Length:          1260
Score:                    472.00        Matches:         91
Percent Similarity:       72.83%        Conservative:    35
Best Local Similarity:    52.60%        Mismatches:      47
Query Match:              53.33%        Indels:          0
DB:                       4             Gaps:            0

US-09-770-509-2 (1-178) x US-09-120-426-1 (1-1260)

Qy          6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
              ||||||||||||||||   ||||||||||||   |||||||||      :::   |||
Db        112 GGCGTAGAATTTATCGCAGCAAACACAGATGTACAAGCATTGAGTAGTACAAAAGCTGAG 171

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
                 |||   ||||||     :::|||:::||||||||||||||||   :::||||||
Db        172 ACTGTTATTCAGTTGGGACCTAAATTGACTCGTGGTTTGGGTGCAGGAGGTCAACCTGAG 231

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
              :::|||::::::::::||||||:::   :::      :::         :::      :::
Db        232 GTTGGTCGTAAAGCCGCTGAAGAAAGCGAAGAAACACTGACGGAAGCTATTAGTGGTGCC 291

Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
              :::|||:::||||||||   ||||||||||||:::   ||||||||||||||||||
Db        292 GATATGGTCTTCATCACTGCTGGTATGGGAGGAGGCTCTGGAACTGGAGCTGCTCCTGTT 351

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
              :::|||   :::|||:::::|||||   ||||||||||||||||:::   ||||||
Db        352 ATTGCTCGTATCGCCAAAGATTTAGGTGCGCTTACAGTTGGTGTTGTAACACGTCCCTTT 411

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
                 ||||||      |||   :::   |||      |||:::   :::|||   :::
Db        412 GGTTTTGAAGGAAGTAAGCGTGGACAATTTGCTGTAGAAGGAATCAATCAACTTCGTGAG 471

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
              :::||||||||||||:::::::::   |||   |||||||||   :::   ||||||
Db        472 CATGTAGACACTCTATTGATTATCTCAAACAACAATTTGCTTGAAATTGTTGATAAGAAA 531

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
              |||   :::||||||||||         ||||||:::||||||   :::|||||||:::|||
Db        532 ACACCGCTTTTGGAGGCTCTTAGCGAAGCGGATAACGTTCTTCGTCAAGGTGTTCAAGGG 591

Qy        166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::||||||||||||      ||||||||||||||||||
Db        592 ATTACCGATTTGATTACCAATCCAGGATTGATTAACCTT 630

RESULT 7
US-08-987-146-4
; Sequence 4, Application US/08987146
; Patent No. 6350866
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Rockey, Pamela K.
;    APPLICANT:  Wang, Q. May
```

```
;    APPLICANT:  Rosteck Jr., Paul R.
;    TITLE OF INVENTION:  Streptococcus Pneumoniae Gene Sequence
;    TITLE OF INVENTION:  FtsZ
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/987,146
;      FILING DATE:  December 8, 1997
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11755
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2702 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-987-146-4

Alignment Scores:
Pred. No.:              8.07e-50      Length:          2702
Score:                  472.00        Matches:         91
Percent Similarity:     72.83%        Conservative:    35
Best Local Similarity:  52.60%        Mismatches:      47
Query Match:            53.33%        Indels:          0
DB:                     4             Gaps:            0

US-09-770-509-2 (1-178) x US-08-987-146-4 (1-2702)

Qy          6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
              ||||||||||||||   ||||||||||   ||||||||      :::   |||
Db        170 GGCGTAGAATTTATCGCAGCAAACACAGATGTACAAGCATTGAGTAGTACAAAAGCTGAG 229

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
                 |||   ||||||      :::|||:::|||||||||||||||   :::|||||||
Db        230 ACTGTTATTCAGTTGGGACCTAAATTGACTCGTGGTTTGGGTGCAGGAGGTCAACCTGAG 289

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
              :::|||::::::::::||||||:::   :::      :::         :::      :::
Db        290 GTTGGTCGTAAAGCCGCTGAAGAAAGCGAAGAAACACTGACGGAAGCTATTAGTGGTGCC 349

Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
              :::|||:::|||||||   ||||||||||||||:::   ||||||||||||||||||
Db        350 GATATGGTCTTCATCACTGCTGGTATGGGAGGAGGCTCTGGAACTGGAGCTGCTCCTGTT 409

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
              :::|||   :::|||:::::::|||||   ||||||||||||||||:::   ||||||
Db        410 ATTGCTCGTATCGCCAAAGATTTAGGTGCGCTTACAGTTGGTGTTGTAACACGTCCCTTT 469

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
                |||||      |||   :::   |||      |||:::   :::|||   :::
Db        470 GGTTTTGAAGGAAGTAAGCGTGGACAATTTGCTGTAGAAGGAATCAATCAACTTCGTGAG 529

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
              :::||||||||||||:::::::::   |||   |||||||||   :::   ||||||

Db        530 CATGTAGACACTCTATTGATTATCTCAAACAACAATTTGCTTGAAATTGTTGATAAGAAA 589

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
                |||   :::||||||||||         ||||||:::||||||   :::|||||:::|||
Db        590 ACACCGCTTTTGGAGGCTCTTAGCGAAGCGGATAACGTTCTTCGTCAAGGTGTTCAAGGG 649

Qy        166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::||||||||||||      ||||||||||||||||||
Db        650 ATTACCGATTTGATTACCAATCCAGGATTGATTAACCTT 688

RESULT 8
US-08-961-527-36
; Sequence 36, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21706 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-527-36

Alignment Scores:
Pred. No.:              1.75e-48      Length:          21706
Score:                  472.00        Matches:         91
Percent Similarity:     72.83%        Conservative:    35
Best Local Similarity:  52.60%        Mismatches:      47
Query Match:            53.33%        Indels:          0
DB:                     4             Gaps:            0

US-09-770-509-2 (1-178) x US-08-961-527-36 (1-21706)

Qy          6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
              ||||||||||||||   ||||||||||   |||||||||      :::   |||
Db       2761 GGCGTAGAATTTATCGCAGCAAACACAGATGTACAAGCATTGAGTAGTACAAAAGCTGAG 2820

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
                 |||   ||||||      :::|||:::|||||||||||||||   :::|||||||
Db       2821 ACTGTTATTCAGTTGGGACCTAAATTGACTCGTGGTTTGGGTGCAGGAGGTCAACCTGAG 2880

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
              :::|||::::::::::||||||:::   :::      :::         :::      :::
Db       2881 GTTGGTCGTAAAGCCGCTGAAGAAAGCGAAGAAACACTGACGGAAGCTATTAGTGGTGCC 2940
```

```
Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
              :::|||:::|||||||||   |||||||||||||||:::   ||||||||||||||||||
Db       2941 GATATGGTCTTCATCACTGCTGGTATGGGAGGAGGCTCTGGAACTGGAGCTGCTCCTGTT 3000

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
              :::|||   :::|||:::::::|||||   ||||||||||||||||||:::   ||||||
Db       3001 ATTGCTCGTATCGCCAAAGATTTAGGTGCGCTTACAGTTGGTGTTGTAACACGTCCCTTT 3060

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
                  ||||||      |||   :::   |||      |||:::   :::|||   :::
Db       3061 GGTTTTGAAGGAAGTAAGCGTGGACAATTTGCTGTAGAAGGAATCAATCAACTTCGTGAG 3120

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
              :::|||||||||||||:::::::::   |||   |||||||||   :::   ||||||
Db       3121 CATGTAGACACTCTATTGATTATCTCAAACAACAATTTGCTTGAAATTGTTGATAAGAAA 3180

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
              |||   :::|||||||||         ||||||:::||||||   :::|||||||:::|||
Db       3181 ACACCGCTTTTGGAGGCTCTTAGCGAAGCGGATAACGTTCTTCGTCAAGGTGTTCAAGGG 3240

Qy        166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||||||||||||      |||||||||||||||||
Db       3241 ATTACCGATTTGATTACCAATCCAGGATTGATTAACCTT 3279

RESULT 9
US-09-381-862-4
; Sequence 4, Application US/09381862
; Patent No. 6245906
;  GENERAL INFORMATION:
;    APPLICANT:  Ueyama, Hiroshi
;    APPLICANT:  Abe, Kanako
;    APPLICANT:  Keshi, Hiroyuki
;    APPLICANT:  Matsuhisa, Akio
;    TITLE OF INVENTION:  PROBES FOR THE DIAGNOSIS OF INFECTIONS
;    TITLE OF INVENTION:  CAUSED BY STREPTOCOCCUS PYOGENES
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/381,862
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 1997-71077
;      FILING DATE:  25-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP98/01288
;      FILING DATE:  23-MAR-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Cawley, Jr., Thomas A.
;      REGISTRATION NUMBER:  40,944
;      REFERENCE/DOCKET NUMBER:  19036/36274
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7143 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  Streptococcus pyogenes
;      STRAIN:  Clinical Isolate SP-26-36
US-09-381-862-4

Alignment Scores:
Pred. No.:              3.49e-48        Length:          7143
Score:                  464.00          Matches:         93
Percent Similarity:     71.10%          Conservative:    30
Best Local Similarity:  53.76%          Mismatches:      50
Query Match:            52.43%          Indels:          0
DB:                     4               Gaps:            0

US-09-770-509-2 (1-178) x US-09-381-862-4 (1-7143)

Qy          6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
              ||||||||||||||   ||||||||||   |||||||||      |||   |||
Db       3673 GGTGTCGAGTTCATCGCAGCAAATACAGACATTCAGGCATTAAGCTCATCAAAAGCTGAA 3732

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
                 |||   ||||||      :::|||:::|||||||||||||||   :::|||||||
Db       3733 ACGGTTATTCAACTAGGCCCTAAATTAACTCGTGGACTTGGTGCTGGAGGACAACCTGAA 3792

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
              :::|||:::::::::|||||:::   :::      :::         |||      :::
Db       3793 GTAGGACGTAAAGCTGCTGAAGAAAGCGAAGAAATTTTAACAGAAGCTCTTACAGGAGCG 3852

Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
              :::|||:::|||||||||   |||||||||||||||:::   ||||||||||||||||||
Db       3853 GATATGGTATTTATTACTGCCGGTATGGGTGGTGGCTCTGGGACAGGGGCTGCACCGGTT 3912

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
              :::|||   :::|||:::   ||||||   |||||||||   |||||||::   ||||||
Db       3913 ATTGCTCGTATCGCTAAGAGTTTGGGAGCCTTGACAGTAGCTGTTGTTACTCGCCCGTTT 3972

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
                  ||||||      |||         |||      |||::::::||||||   :::
Db       3973 GGTTTTGAAGGTAACAAACGTGGTAATTTTGCTATTGAAGGTATCGAAGAACTCCGTGAA 4032

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
              ||||||||||||:::::::::   |||   |||||||||   :::   ||||||
Db       4033 CAAGTTGATACTTTGTTAATTATTTCAAATAATAACCTTCTTGAGATTGTTGATAAAAAG 4092

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
              |||   :::|||||||||         ||||||:::||||||   :::|||||||:::|||
Db       4093 ACACCTTTATTAGAAGCACTTAGTGAAGCTGATAATGTTTTACGTCAGGGAGTTCAAGGG 4152

Qy        166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||||||||||||      |||||||||||||||||
Db       4153 ATAACCGACTTAATTACTAGTCCTGGCCTTATCAATCTC 4191

RESULT 10
US-08-883-515-3
; Sequence 3, Application US/08883515
; Patent No. 5981836
;  GENERAL INFORMATION:
;    APPLICANT:  Osteryoung, Katherine W
;    TITLE OF INVENTION:  PLANT CHLOROPLAST DIVISION GENES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  1 South Pinckney Street
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  US
;      ZIP:  53701-2113
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,515
```

```
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seay, Nicholas J.
;      REGISTRATION NUMBER:  27,386
;      REFERENCE/DOCKET NUMBER:  920905.90016
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1628 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..1316
US-08-883-515-3

Alignment Scores:
Pred. No.:              5.26e-49      Length:        1628
Score:                  463.00        Matches:       89
Percent Similarity:     69.83%        Conservative:  36
Best Local Similarity:  49.72%        Mismatches:    52
Query Match:            52.32%        Indels:        2
DB:                     2             Gaps:          1

US-09-770-509-2 (1-178) x US-08-883-515-3 (1-1628)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||::::::   |||||||||||||   :::   |||||||||   |||||||:::
Db        297 AGTGAAATGTCAGGTGTGGAGTTCTGGATTGTCAACACTGATATCCAGGCTATGAGAATG 356

Qy         22 Ser------LeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGly 39
              |||      |||      :::::::::   :::||||||::::::|||:::|||||||||
Db        357 TCTCCTGTTTTGCCTGATAATAGGTTACAAATTGGTAAGGAGTTGACTAGGGGTTTAGGT 416

Qy         40 AlaGlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGln 59
              ||||||      ||||||:::|||      :::|||   :::   |||      |||:::
Db        417 GCTGGAGGAAATCCAGAAATCGGTATGAATGCTGCTAGAGAGAGCAAAGAAGTTATTGAA 476

Qy         60 ArgMetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCys 79
                 |||      |||:::|||:::|||:::|||  .||||||||||||||||||
Db        477 GAAGCTCTTTATGGCTCAGATATGGTCTTTGTCACAGCTGGAATGGGCGGTGGAACTGGC 536

Qy         80 ThrGlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGly 99
              ||||||||||||||||:::|||   :::|||:::   :::|||||||||||||||||
Db        537 ACTGGTGCAGCCCCTGTAATTGCAGGAATTGCCAAGGCGATGGGTATATTGACAGTTGGT 596

Qy        100 ValValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGly 119
              :::   :::|||||||||      ||||||      ||||||      |||      |||
Db        597 ATTGCCACAACGCCTTTCTCGTTTGAGGGTCGAAGAAGAACTGTTCAGGCTCAAGAAGGG 656

Qy        120 ValLysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeu 139
              :::      |||         ||||||||||||||||||:::||||||      ||||||
Db        657 CTTGCATCTCTCAGAGACAATGTTGACACTCTCATCGTCATTCCAAATGACAAGTTGCTT 716

Qy        140 AlaLeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeu 159
                    :::||||||   :::   |||||||||      |||||||||:::|||
Db        717 ACAGCTGTCTCTCAGTCTACTCCGGTAACAGAAGCATTTAATCTAGCTGATGATATACTC 776

Qy        160 LeuGluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
               :::||||||:::|||::::::|||:::|||      |||||||||:::|||:::
Db        777 CGTCAGGGGGTTCGTGGGATATCTGATATCATTACGATTCCTGGTTTGGTGAATGTG 833

RESULT 11
US-08-883-515-1
;  GENERAL INFORMATION:
;    APPLICANT:  Austin-Phillips, Sandra
;    APPLICANT:  Burgess, Richard R.
;    APPLICANT:  German, Thomas L.
;    APPLICANT:  Ziegelhoffer, Thomas L.
;    TITLE OF INVENTION:  Transgenic Plants as an Alternative
;    TITLE OF INVENTION:  Source of Lignocellulosic-Degrading Enzymes
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DeWitt Ross & Stevens, S.C.
;      STREET:  8000 Excelsior Drive, Ste. 401
;      CITY:  Madison
;      STATE:  Wisconsin
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,495
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/028,718
;      FILING DATE:  17-OCT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sara, Charles S.
;      REGISTRATION NUMBER:  30,492
;      REFERENCE/DOCKET NUMBER:  09820.036
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
; Sequence 1, Application US/08883515
; Patent No. 5981836
;  GENERAL INFORMATION:
;    APPLICANT:  Osteryoung, Katherine W
;    TITLE OF INVENTION:  PLANT CHLOROPLAST DIVISION GENES
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  1 South Pinckney Street
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  US
;      ZIP:  53701-2113
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/883,515
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seay, Nicholas J.
;      REGISTRATION NUMBER:  27,386
;      REFERENCE/DOCKET NUMBER:  920905.90016
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1425 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  25..1326
US-08-883-515-1

Alignment Scores:
Pred. No.:              8.68e-45      Length:        1425
```

Score:	429.00	Matches:	86
Percent Similarity:	70.79%	Conservative:	40
Best Local Similarity:	48.31%	Mismatches:	52
Query Match:	48.47%	Indels:	0
DB:	2	Gaps:	0

US-09-770-509-2 (1-178) x US-08-883-515-1 (1-1425)

```
Qy          1 AlaSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
              :::|||   |||:::   |||:::|||         |||||||||   |||||||||
Db        304 TCAAGCGGTTTACAGAGTGTTGATTTCTATGCGATAAACACGGATTCGCAAGCTCTGTTA 363

Qy         21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
              :::      |||   :::   :::   :::|||:::   :::|||:::|||||||||
Db        364 CAGTTTTCTGCTGAGAACCCACTTCAAATTGGAGAACTTTTAACTCGTGGGCTTGGCACT 423

Qy         41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
              |||      |||   ||||||::::::::::||||||:::   |||      |||
Db        424 GGTGGAAACCCGCTTCTTGGAGAACAAGCTGCAGAAGAATCAAAAGATGCAATTGCTAAT 483

Qy         61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
                 |||:::   |||:::::::::|||||||||   ||||||||||||||||||   :::
Db        484 GCTCTTAAAGGATCAGACCTTGTTTTCATAACTGCTGGTATGGGTGGTGGAACAGGGTCT 543

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
              ||||||||||||||||||||   ::::::::::::   |||   ||||||||||||||||
Db        544 GGTGCTGCACCTGTGGTAGCTCAGATTTCGAAGGATGCTGGTTATTTGACTGTTGGTGTT 603

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
              |||:::   ||||||      ||||||      |||:::      |||         :::
Db        604 GTTACCTATCCGTTTAGCTTTGAAGGACGTAAAAGATCTTTGCAGGCACTGGAAGCTATT 663

Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
              :::::::|||   |||   ||||||||||||||||||:::||||||      ||||||
Db        664 GAAAAGCTCCAAAAGAATGTTGATACCCTTATCGTGATTCCAAATGATCGTCTGCTAGAT 723

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
              :::|||||||:::   |||   :::   :::||||||      .||||||||||||||||
Db        724 ATTGCTGATGAACAGACGCCACTTCAGGACGCGTTTCTTCTTGCAGATGATGTTTTACGC 783

Qy        161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::||||||:::|||:::::|||:::|||      |||||||||:::|||:::
Db        784 CAAGGAGTACAAGGAATCTCAGATATTATTACTATACCTGGACTAGTCAATGTG 837
```

RESULT 12
US-09-134-001C-1996/c
; Sequence 1996, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1996
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1996

Alignment Scores:			
Pred. No.:	6.68e-39	Length:	396
Score:	376.00	Matches:	72
Percent Similarity:	76.15%	Conservative:	27
Best Local Similarity:	55.38%	Mismatches:	31
Query Match:	42.49%	Indels:	0
DB:	4	Gaps:	0

US-09-770-509-2 (1-178) x US-09-134-001C-1996 (1-396)

```
Qy         30 LeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeuGlyLysArg 49
              :::|||:::   :::|||:::|||||||||||||||:::   ||||||:::||||||:::
Db        392 ATTGGTGAGAAATTAACACGTGGATTAGGTGCTGGTGCTAATCCTGAAATTGGGAAGAAA 333

Qy         50 SerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsnMetLeuPhe 69
              :::||||||:::   :::      |||:::      :::|||   ::::::|||:::|||
Db        332 GCAGCAGAAGAATCAAGAGAACAAATTGAAGACGCTATCCAAGGTGCTGATATGGTATTC 273

Qy         70 IleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValValAlaSerVal 89
              :::|||   ||||||||||||||||||   ||||||||||||||||||||||||   :::
Db        272 GTAACTGCTGGTATGGGTGGCGGTACTGGAACAGGTGCTGCACCGGTTGTTGCTAAAATA 213

Qy         90 AlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArgSerGluGly 109
              |||:::|||:::|||   ||||||||||||||||||:::   ||||||      ||||||
Db        212 GCAAAAGAAATGGGTGCTTTAACTGTAGGTGTTGTTACGCGTCCATTCGGTTTCGAAGGT 153

Qy        110 ProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyrValAspThr 129
                    |||         |||   |||||||||:::   :::         ||||||||
Db        152 CGTAAGCGCCAAACACAAGCGGCAGCTGGCGTAGAATCTATGAAAGCAGCAGTGGATACA 93

Qy        130 LeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThrThrMetLeu 149
              |||||||||:::||||||      ||||||   :::   ||||||||||   |||:::
Db         92 TTAATTGTTATTCCAAATGATCGCTTATTAGATATCGTTGACAAATCTACGCCAATGATG 33

Qy        150 GluAlaPheArgTyrAlaAspAspValLeu 159
              |||||||||:::   |||||||:::||||||
Db         32 GAAGCATTTAAAGAAGCGGATAATGTATTA 3
```

RESULT 13
US-09-221-017B-658
; Sequence 658, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600

```
;      TELEFAX:  650-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  658:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  665 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  UNKNOWN
;    ORIGINAL SOURCE:
;      ORGANISM:  PORYPHYROMONAS GINGIVALIS
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1...665
US-09-221-017B-658

Alignment Scores:
Pred. No.:              9.55e-38        Length:         665
Score:                  369.50          Matches:        76
Percent Similarity:     67.92%          Conservative:   32
Best Local Similarity:  47.80%          Mismatches:     48
Query Match:            41.75%          Indels:         3
DB:                     4               Gaps:           2

US-09-770-509-2 (1-178) x US-09-221-017B-658 (1-665)

Qy          3 GlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSer 22
              ::::::      |||   |||::::::   |||||||||   |||||||||   ||||||
Db        183 AAGGTGCGGGATGTATCGTTCCTCCTGTGTAATACCGATGTGCAGGCTCTCGACCGAAGC 242

Qy         23 LeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySer 42
                 |||   ::::::   ||||||:::::::::|||   ||||||||||||||||||
Db        243 GAGGTACCTGATCGGCTGGTACTCGGCCGTGAGGTGACCAACGGTCTGGGTGCCGGTAGC 302

Qy         43 LysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeu 62
              :::||||||:::   :::|||:::||||||      :::   ||||||:::::::::|||
Db        303 CGTCCGGAAGTGGCACGACGAGCTGCAGAGGCCAGCGAAGCCGACATACGCAAGATATTG 362

Qy         63 GlnAsp------SerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
                 |||      :::   |||:::|||:::|||   ||||||||||||||||||   |||
Db        363 GATGATGGCCATACCCGTATGGTCTTCGTAACGGCCGGTATGGGCGGCGGAACGGGTACC 422

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
              ||||||||||||||:::      :::||||||||||||   ||||||||||||||:::
Db        423 GGTGCGGCCCCCTGTCATCGGTCGGATAGCACGTGAACTCAATATCCTCACTGTCGGCATC 482

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
              |||:::   ||||||      ||||||      :::         |||      ||||||
Db        483 GTTACCATCCCATTCGTTTTCGAAGGCAAGCGCAAGATTCTGCAGGCACTCGAAGGGGTG 542

Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
              :::|||:::   |||   ||||||   |||:::||||||   |||:::      |||
Db        543 GAGGAAATGCGCAAAAACGTAGATGCCTTGCTCGTGGTCAATAATGAACGG---CTCCGC 599

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeu 159
              :::      |||         :::      ||||||      ||||||:::   |||
Db        600 ATCATATACAAAGATCTTAAGCTCGACAATGCTTTTGCCAAGGCAGACGAGACACTG 656

RESULT 14
US-09-120-426-3
; Sequence 3, Application US/09120426
; Patent No. 6197300
; GENERAL INFORMATION:
;  APPLICANT: Fueyo, Joanna Lynn
;  APPLICANT: Lonetto, Michael A.
;  APPLICANT: Pearce, Kenneth
;  TITLE OF INVENTION: ftsZ
;  FILE REFERENCE: GM10068
;  CURRENT APPLICATION NUMBER: US/09/120,426
;  CURRENT FILING DATE: 1998-07-22
;  EARLIER APPLICATION NUMBER: 60/055,720
;  EARLIER FILING DATE: 1997-08-12
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 669
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(669)
US-09-120-426-3

Alignment Scores:
Pred. No.:              2.82e-25        Length:         669
Score:                  271.00          Matches:        52
Percent Similarity:     74.23%          Conservative:   20
Best Local Similarity:  53.61%          Mismatches:     25
Query Match:            30.62%          Indels:         0
DB:                     4               Gaps:           0

US-09-770-509-2 (1-178) x US-09-120-426-3 (1-669)

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              :::||||||||||:::|||   :::|||:::::::||||||   ||||||||||||||||||
Db          1 TCTGCTCCTGTTATTGCTCGTATCGCCAAAGATTTAGGTGCGCTTACAGTTGGTGTTGTA 60

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||      |||   :::   |||      |||:::
Db         61 ACACGTCCCTTTGGTTTTGAAGGAAGTAAGCGTGGACAATTTGCTGTAGAAGGAATCGAT 120

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              ::::||   ::::::||||||||||||||||:::::::   |||   |||||||||   :::
Db        121 CAACTTCGTGAGCATGTAGACACTCTATTGATTATCTCAAACAACAATTTGCTTGAAATT 180

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
                 ||||||   |||   :::|||||||||         ||||||:::||||||   :::
Db        181 GTTGATAAGAAAACACCGCTTTTGGAGGCTCTTAGCGAAGCGGATAACGTTCTTCGTCAA 240

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||:::||||||||||||      ||||||||||||||||||
Db        241 GGTGTTCAAGGGATTACCGATTTGATTACCAATCCAGGATTGATTAACCTT 291

RESULT 15
US-08-961-083-123
; Sequence 123, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
;   APPLICANT:  Choi et. al.
;   TITLE OF INVENTION:  Streptococcus pneumoniae Antigens and Vaccines
;   NUMBER OF SEQUENCES:  452
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Human Genome Sciences, Inc.
;     STREET:  9410 Key West Avenue
;     CITY:  Rockville
;     STATE:  Maryland
;     COUNTRY:  USA
;     ZIP:  20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;     COMPUTER:  HP Vectra 486/33
;     OPERATING SYSTEM:  MSDOS version 6.2
;     SOFTWARE:  ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/961,083
;     FILING DATE:
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Brookes, A. Anders
;     REGISTRATION NUMBER:  36,373
;     REFERENCE/DOCKET NUMBER:  PB340P2
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  123:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  850 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-083-123

Alignment Scores:
Pred. No.:              1.25e-15        Length:         850
Score:                  196.00          Matches:        38
Percent Similarity:     70.27%          Conservative:   14
Best Local Similarity:  51.35%          Mismatches:     22
Query Match:            22.15%          Indels:         0
DB:                     3               Gaps:           0

US-09-770-509-2 (1-178) x US-08-961-083-123 (1-850)

Qy        105 PheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAla 124
              |||       ||||||       |||   :::   |||       |||:::   :::|||
Db          2 TTTGGTTTTGAAGGAAGTAAGCGTGGACAATTTGCTGTAGAAGGAATCAATCAACTTCGT 61

Qy        125 LysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLys 144
              ::::::||||||||||||:::::::::   |||   |||||||||   :::   ||||||
Db         62 GAGCATGTAGACACTCTATTGATTATCTCAAACAACAATTTGCTTGAAATTGTTGATAAG 121

Qy        145 SerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLys 164
                 |||   :::|||||||||         ||||||:::||||||   :::||||||:::
Db        122 AAAACACCGCTTTTGGAGGCTCTTAGCGAAGCGGATAACGTTCTTCGTCAAGGTGTTCAA 181

Qy        165 GlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              |||:::|||||||||||      ||||||||||||||||||
Db        182 GGGATTACCGATTTGATTACCAATCCAGGATTGATTAACCTT 223


Search completed: June  2, 2003, 14:09:19
Job time : 14522 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 10:13:26 ; Search time 170 Seconds
(without alignments)
1411.818 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 845702 seqs, 674182571 residues

Total number of hits satisfying chosen parameters: 1691404

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09770509/runat_27052003_083654_19004/app_query.fasta_1.327
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09770509_@CGN_1_1_138_@runat_27052003_083654_19004
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT
-DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10
-XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	885	100.0	535	9	US-09-770-509-1	Sequence 1, Appli
2	509	57.5	950	10	US-09-974-300-4664	Sequence 4664, Ap
3	502	56.7	1134	10	US-09-974-300-180	Sequence 180, App
4	495	55.9	1233	10	US-09-815-242-6542	Sequence 6542, Ap
5	495	55.9	1239	10	US-09-815-242-3854	Sequence 3854, Ap
c 6	495	55.9	32768	10	US-09-070-927A-128	Sequence 128, App
7	485	54.8	1140	9	US-09-712-363-75	Sequence 75, Appl
8	481	54.4	1423	9	US-09-770-509-9	Sequence 9, Appli
9	480	54.2	1172	10	US-09-815-242-4715	Sequence 4715, Ap
10	480	54.2	1179	10	US-09-815-242-8625	Sequence 8625, Ap
11	480	54.2	3931	7	US-08-781-986A-342	Sequence 342, App
12	472	53.3	1260	10	US-09-754-608-1	Sequence 1, Appli
13	472	53.3	1260	10	US-09-815-242-9381	Sequence 9381, Ap
14	470	53.1	1326	9	US-09-738-626-2360	Sequence 2360, Ap
c 15	470	53.1	3309400	9	US-09-738-626-1	Sequence 1, Appli
16	464	52.4	1185	10	US-09-815-242-7909	Sequence 7909, Ap
c 17	461	52.1	640681	10	US-09-790-988-1	Sequence 1, Appli
18	452	51.1	1152	10	US-09-912-020-105	Sequence 105, App
19	452	51.1	1152	10	US-09-815-242-5936	Sequence 5936, Ap
20	447	50.5	1152	9	US-09-908-504A-2	Sequence 2, Appli
21	447	50.5	1266	10	US-09-815-242-7095	Sequence 7095, Ap
c 22	443	50.1	607	10	US-09-815-242-1875	Sequence 1875, Ap
c 23	443	50.1	607	10	US-09-815-242-2202	Sequence 2202, Ap
c 24	443	50.1	607	10	US-09-815-242-2203	Sequence 2203, Ap
25	435	49.2	1302	9	US-09-938-842A-1876	Sequence 1876, Ap
26	379	42.8	1158	10	US-09-815-242-7456	Sequence 7456, Ap
c 27	347	39.2	525	10	US-09-815-242-289	Sequence 289, App
c 28	313	35.4	360	10	US-09-815-242-774	Sequence 774, App
c 29	313	35.4	362	10	US-09-815-242-854	Sequence 854, App
c 30	311	35.1	360	10	US-09-815-242-795	Sequence 795, App
31	271	30.6	669	10	US-09-754-608-3	Sequence 3, Appli
c 32	235	26.6	231	10	US-09-815-242-640	Sequence 640, App
33	196	22.1	850	10	US-09-765-272-123	Sequence 123, App
34	160	18.1	220	9	US-09-770-509-3	Sequence 3, Appli
c 35	132	14.9	111	10	US-09-815-242-892	Sequence 892, App
c 36	132	14.9	111	10	US-09-815-242-1038	Sequence 1038, Ap
37	114	12.9	388	9	US-09-770-509-5	Sequence 5, Appli
c 38	109	12.3	200	10	US-09-815-242-822	Sequence 822, App
c 39	108	12.2	266	10	US-09-815-242-1818	Sequence 1818, Ap
c 40	99	11.2	204	10	US-09-815-242-751	Sequence 751, App
41	98	11.1	205	10	US-09-770-696-135	Sequence 135, App
c 42	91	10.3	87	10	US-09-815-242-3504	Sequence 3504, Ap
c 43	91	10.3	87	10	US-09-815-242-3598	Sequence 3598, Ap
c 44	91	10.3	87	10	US-09-815-242-3614	Sequence 3614, Ap
45	90.5	10.2	7100	10	US-09-932-183A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-770-509-1
; Sequence 1, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
; APPLICANT: Katagiri, F.
; TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
; TITLE OF INVENTION: ANTIMICROBIAL-SPECIFIC BIOCIDES
; FILE REFERENCE: NADII.018A
; CURRENT APPLICATION NUMBER: US/09/770,509
; CURRENT FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Phytophthora infestans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(535)
US-09-770-509-1

Alignment Scores:

Pred. No.:	2.03e-107	Length:	535
Score:	885.00	Matches:	178
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

```
DB:                    9              Gaps:          0

US-09-770-509-2 (1-178) x US-09-770-509-1 (1-535)

Qy          1 AlaSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCGTCGCAATTGGAAGGTGTGGAGTTCATTGTAGCCAACACAGACTGTCAGGCTCTGGGA 61

Qy         21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CGCTCGCTGGCGCCGCACAAGATCACGCTGGGCAAAGATATCACCAAGGGACTAGGAGCT 121

Qy         41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGATCCAAACCTGAGCTGGGTAAACGCTCTGCGGAACAGCAGAAAGTGGATATCCAACGG 181

Qy         61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATGTTACAGGACAGCAACATGCTGTTTATCACGGGCGGAATGGGCGGCGGAACCTGCACA 241

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGAGCCGCACCTGTCGTGGCCAGTGTAGCCAGGGAGCTGGGGATCCTAACGGTCGGAGTA 301

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GTAAGCACACCGTTCCGATCCGAAGGACCCAATCGCACTCGTCTGGCCAATGCTGGAGTA 361

Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AAAGAACTGGCCAAGTACGTCGACACCTTAATTGTCGTGCCCAACCAGAACTTGCTGGCT 421

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTGGCAGACAAGAGCACGACCATGTTGGAAGCCTTCCGGTATGCCGACGACGTGCTGCTT 481

Qy        161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAAGGAGTTAAAGGTGTCACGGACTTGATCGTTCGCCCGGGACTTATCAATTTG 535

RESULT 2
US-09-974-300-4664
; Sequence 4664, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4664
;   LENGTH: 950
;   TYPE: DNA
;   ORGANISM: Bacillus clausii
US-09-974-300-4664

Alignment Scores:
Pred. No.:              1.57e-57        Length:         950
Score:                  509.00          Matches:        101
Percent Similarity:     74.86%          Conservative:   30
Best Local Similarity:  57.71%          Mismatches:     44
Query Match:            57.51%          Indels:         0
DB:                     10              Gaps:           0

US-09-770-509-2 (1-178) x US-09-974-300-4664 (1-950)
```

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||:::|||||      |||||||||   |||||||||      |||
Db        103 CTGCAAGGTGTTGATTTTATTGCAGTCAATACAGATGCACAAGCACTCCATTTATCAAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   ||||||      :::|||:::||||||||||||||||:::
Db        163 GCGGAAAAGAAGTTGCAACTTGGCGGCAAACTTACTCGCGGCCTTGGCGCTGGCGCCAAC 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||::::::|||||||:::   :::      ::::::   :::|||
Db        223 CCAGAAATAGGCAAGAAAGCGGCAGAGGAAAGCAGAGAACAGCTGGAAGAAGTTCTTACA 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||:::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        283 GGTTCTGATATGGTCTTTATTACTGCCGGAATGGGAGGAGGAACAGGCACCGGAGCAGCT 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   ||||||:::|||||||||   ||||||||||||||||||:::
Db        343 CCTGTCATTGCTGAAGTGGCAAAAGAGCTTGGCGCATTGACGGTCGGTGTTGTGACACGC 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||   :::|||:::      |||
Db        403 CCTTTCTCTTTTGAAGGCCGTAAGCGGCAAAACCAAGCGATCTCTGGGATTGCAGCGTTA 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||||||||:::||||||      ||||||   :::   |||
Db        463 AAAGAAAAAGTGGACACATTGATCGTGATTCCAAATGACCGCTTATTGGAAATGGTTGAT 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   ||||||||||||||||||   ||||||:::||||||   :::|||||||
Db        523 AAAAATACGCCGATGCTTGAAGCGTTCCGCGAGGCGGATAATGTATTGCGCCAAGGAGTT 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||||||||||||
Db        583 CAAGGCATTTCCGATTTAATTGCCACGCCTGGGTTGATCAACCTT 627

RESULT 3
US-09-974-300-180
; Sequence 180, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 180
;   LENGTH: 1134
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-180

Alignment Scores:
Pred. No.:              1.7e-56         Length:         1134
Score:                  502.00          Matches:        97
Percent Similarity:     74.01%          Conservative:   34
Best Local Similarity:  54.80%          Mismatches:     46
Query Match:            56.72%          Indels:         0
DB:                     10              Gaps:           0

US-09-770-509-2 (1-178) x US-09-974-300-180 (1-1134)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              :::   ::::::|||||||||||||||      |||||||||   |||||||||
```

```
Db         97 AATGACGTTCAGGGAGTCGAGTTTATCGCAGTCAACACGGATGCTCAGGCTCTCAACCTG 156

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||   |||      |||:::   :::|||      :::|||:::|||||||||||||||
Db        157 TCAAAAGCGGAAACGAAAATGCAGATCGGTGCGAAGCTGACGCGCGGGCTCGGCGCCGGA 216

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   ||||||:::||||||:::::||||||:::   |||      |||:::
Db        217 GCCAATCCGGAAGTGGGCAAAAAAGCCGCAGAGGAAAGCAAAGAACAAATTGAAGAAGCG 276

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::   ::::::|||:::|||:::|||   ||||||||||||||||||   ||||||
Db        277 CTGAAAGGTGCAGATATGGTGTTCGTCACAGCCGGAATGGGCGGCGGAACAGGAACGGGC 336

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||:::|||   :::|||:::::||||||   ||||||||||||||||||
Db        337 GCGGCACCTGTCATCGCACAAATCGCAAAAGATCTGGGCGCATTGACTGTCGGCGTTGTC 396

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||      |||          |||      |||:::
Db        397 ACAAGGCCGTTTACCTTTGAAGGAAGAAAAAGACAGCTTCAGGCTGCAGGCGGTATTTCA 456

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
                 :::   :::   ||||||||||||||||||:::||||||      ||||||   :::
Db        457 GCAATGAAGGAAGCCGTCGACACCCTGATCGTCATTCCGAATGACCGTCTTCTTGAAATC 516

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              |||||:::|||   ||||||||||||||||||   ||||||:::||||||   :::
Db        517 GTCGATAAAAACACACCGATGCTTGAAGCGTTCCGTGAAGCGGACAACGTTCTCCGCCAA 576

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||:::::|||||||||      ||||||||||||||||||
Db        577 GGTGTTCAGGGCATTTCAGACCTGATCGCAACGCCTGGACTGATCAACCTT 627

RESULT 4
US-09-815-242-6542
; Sequence 6542, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6542
;   LENGTH: 1233
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1233)
US-09-815-242-6542

Alignment Scores:
Pred. No.:                1.6e-55        Length:          1233
Score:                    495.00         Matches:         100
Percent Similarity:       73.14%         Conservative:    28
Best Local Similarity:    57.14%         Mismatches:      47
Query Match:              55.93%         Indels:          0
DB:                       10             Gaps:            0

US-09-770-509-2 (1-178) x US-09-815-242-6542 (1-1233)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||   ||||||||||   ||||||||      |||
Db        106 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 165

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||         |||:::||||||||||||||||||:::
Db        166 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 225

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::||||||:::   :::      |||         ||||||
Db        226 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 285

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        286 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 345

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db        346 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 405

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||      |||:::      |||
Db        406 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCCGCTGAAGGAATTGCCTTATTA 465

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::   ||||||||||||:::::::::   |||      ||||||   :::   |||
Db        466 AAAGAAAACGTTGATACACTATTAATTATCTCAAACAACCGCTTATTAGAAGTCGTTGAC 525

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db        526 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 585

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::|||||||||      ||||||   :::||||||
Db        586 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTA 630

RESULT 5
US-09-815-242-3854
; Sequence 3854, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
```

```
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3854
;   LENGTH: 1239
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
US-09-815-242-3854

Alignment Scores:
Pred. No.:                1.61e-55        Length:          1239
Score:                    495.00          Matches:         100
Percent Similarity:       73.14%          Conservative:    28
Best Local Similarity:    57.14%          Mismatches:      47
Query Match:              55.93%          Indels:          0
DB:                       10              Gaps:            0

US-09-770-509-2 (1-178) x US-09-815-242-3854 (1-1239)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||   ||||||||||||   |||||||||      |||
Db        106 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 165

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||        |||:::||||||||||||||||||:::
Db        166 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 225

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||:::   :::      |||         ||||||
Db        226 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 285

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              :::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        286 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 345

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db        346 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 405

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||      |||:::      |||
Db        406 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCCGCTGAAGGAATTGCCTTATTA 465

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::   ||||||||||||:::::::::   |||      ||||||   :::   |||
Db        466 AAAGAAAACGTTGATACACTATTAATTATCTCAAACAACCGCTTATTAGAAGTCGTTGAC 525

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::|||||||
Db        526 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 585

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||   :::||||||
Db        586 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTA 630

RESULT 6
US-09-070-927A-128/c
; Sequence 128, Application US/09070927A
; Patent No. US20020120116A1
;   GENERAL INFORMATION:
;        APPLICANT: Charles A. Kunsch
;                   Patrick J. Dillon
;                   Steven Barash
;        TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
;        NUMBER OF SEQUENCES: 982
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/070,927A
;             FILING DATE: 04-May-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/046,655
;             FILING DATE: 1997-05-16
;             APPLICATION NUMBER: 60/044,031
;             FILING DATE: 1997-05-06
;             APPLICATION NUMBER: 60/066,009
;             FILING DATE: 1997-11-14
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Kenley K. Hoover
;             REGISTRATION NUMBER: 40,302
;             REFERENCE/DOCKET NUMBER: PB369
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 128:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 32768 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 128:
US-09-070-927A-128

Alignment Scores:
Pred. No.:                1.81e-53        Length:          32768
Score:                    495.00          Matches:         100
Percent Similarity:       73.14%          Conservative:    28
Best Local Similarity:    57.14%          Mismatches:      47
Query Match:              55.93%          Indels:          0
DB:                       10              Gaps:            0

US-09-770-509-2 (1-178) x US-09-070-927A-128 (1-32768)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||   ||||||||||||   |||||||||      |||
Db      27971 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 27912

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||        |||:::||||||||||||||||||:::
Db      27911 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 27852

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||:::   :::      |||         ||||||
Db      27851 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 27792

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              :::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db      27791 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 27732

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db      27731 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 27672

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||      |||:::      |||
Db      27671 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTGCNTGCTGAAGGAATTGCCTTATTA 27612
```

```
Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::  ||||||||||||:::::::::  |||      ||||||   :::  |||
Db      27611 AAAGAAAACGTTGATACACTATTAATTATCTCAAATAACCGCTTATTAGAAGTCGTTGAC 27552

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db      27551 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 27492

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::||||||||      ||||||   :::||||||
Db      27491 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTG 27447

RESULT 7
US-09-712-363-75
; Sequence 75, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
;  APPLICANT: Eisenberg, David
;  APPLICANT:  Rotstein, Sergio H.
;  APPLICANT:  Marcotte, Edward M.
;  TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
;  TITLE OF INVENTION:  INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
;  FILE REFERENCE: 07419-032001
;  CURRENT APPLICATION NUMBER: US/09/712,363
;  CURRENT FILING DATE:  2000-11-13
;  PRIOR APPLICATION NUMBER: PCT/US00/02246
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: 60/179,531
;  PRIOR FILING DATE: 2000-02-01
;  PRIOR APPLICATION NUMBER: 60/117,844
;  PRIOR FILING DATE: 1999-01-29
;  PRIOR APPLICATION NUMBER: 60/118,206,
;  PRIOR FILING DATE: 1999-02-01
;  PRIOR APPLICATION NUMBER: 60/126,593
;  PRIOR FILING DATE: 1999-03-26
;  PRIOR APPLICATION NUMBER: 60/134,093
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/134,092
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/165,124
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/165,086
;  PRIOR FILING DATE: 1999-11-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 75
;   LENGTH: 1140
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
US-09-712-363-75

Alignment Scores:
Pred. No.:              2.97e-54        Length:         1140
Score:                  485.00          Matches:        95
Percent Similarity:     74.29%          Conservative:   35
Best Local Similarity:  54.29%          Mismatches:     45
Query Match:            54.80%          Indels:         0
DB:                     9               Gaps:           0

US-09-770-509-2 (1-178) x US-09-712-363-75 (1-1140)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||||||||||      |||||||||   |||||||||      |||
Db         94 CTCAAAGGCGTGGAATTCATCGCGATCAACACCGACGCCCAGGCGTTGTTGATGAGCGAT 153

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||:::|||   |||:::||||||||||||||:::
Db        154 GCCGACGTCAAACTCGACGTCGGCCGCGACTCCACCCGCGGGCTGGGCGCCGGCGCCGAT 213

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||      |||   :::|||:::   :::|||:::
Db        214 CCGGAGGTCGGCCGTAAGGCCGCCGAGGACGCCAAGGACGAGATCGAAGAGCTGCTGCGC 273

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||:::|||   |||   ||||||||||||   ||||||   |||
Db        274 GGTGCCGACATGGTGTTTGTCACCGCCGGCGAGGGGGGCGGAACCGGCACCGGGGGGGCA 333

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              |||||||||||||||:::||||||:::||||||   ||||||||||||||||||:::
Db        334 CCCGTCGTCGCCAGCATCGCCCGCAAGCTGGGCGCGTTGACCGTCGGTGTGGTCACCCGG 393

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||:::      |||      |||:::      |||
Db        394 CCGTTCTCGTTCGAGGGCAAGCGACGCAGCAATCAGGCCGAAAATGGCATCGCGGCGCTG 453

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::      ||||||||||||||||:::||||||      ||||||   :::   |||
Db        454 CGGGAGAGTTGCGACACCCTCATCGTGATTCCCAACGACCGGTTGCTGCAGATGGGAGAT 513

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              :::   ::::::::::::|||||||||   ||||||:::|||||||||   ||||||
Db        514 GCCGCGGTATCGCTGATGGATGCTTTCCGTAGCGCCGACGAGGTGCTGCTCAACGGCGTG 573

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::|||||||||||||      |||||||||||||||:::
Db        574 CAGGGCATCACCGACCTGATTACCACCCCGGGTCTAATCAACGTC 618

RESULT 8
US-09-770-509-9
; Sequence 9, Application US/09770509
; Publication No. US20030082657A1
; GENERAL INFORMATION:
;  APPLICANT: Katagiri, F.
;  TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
;  TITLE OF INVENTION:  ANTIMICROBIAL-SPECIFIC BIOCIDES
;  FILE REFERENCE: NADII.018A
;  CURRENT APPLICATION NUMBER: US/09/770,509
;  CURRENT FILING DATE:  2001-01-25
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 1423
;   TYPE: DNA
;   ORGANISM: Phytophthora infestans
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)...(1261)
US-09-770-509-9

Alignment Scores:
Pred. No.:              1.37e-53        Length:         1423
Score:                  481.00          Matches:        92
Percent Similarity:     73.60%          Conservative:   39
Best Local Similarity:  51.69%          Mismatches:     47
Query Match:            54.35%          Indels:         0
DB:                     9               Gaps:           0

US-09-770-509-2 (1-178) x US-09-770-509-9 (1-1423)

Qy          1 AlaSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
              |||      |||:::|||||||||||||:::|||   |||||||||   |||   |||
Db        386 GCGCGCGGCCTGCAGGGTGTGGAGTTTCTTGTTTGCAACACGGATGCTCAGCACTTACGC 445

Qy         21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
              :::|||      :::::::::   :::      ::::::|||   |||||||||
Db        446 ACGACGCTGACGGAGAACCGCGTTCAGATGGCTCCTGAATTGACTGGAGGACTGGGCTGT 505

Qy         41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
              |||:::   ||||||:::|||:::   :::|||||||            :::|||
Db        506 GGCGCTAACCCCGAAGTTGGCCGAGAGGCGGCAGAGGCCGCGATTGATGAGATTTTGGAG 565

Qy         61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
              :::|||   :::||||||:::|||:::|||   ||||||||||||||||||   |||
Db        566 CGCGTTCAGGGTGCAAACATGATGTTTGTTACTGCGGGTATGGGTGGCGGAACAGGTACA 625

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
```

```
                ||||||||||||||||:::|||      |||   :::   ||||||||||||||||   |||
Db        626 GGTGCAGCACCCGTCATTGCTCAGGCTGCCTTAGATGCTGGTATCCTCACCGTAGCTGTC 685

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
                |||:::   |||||||||   ||||||   ||||||   :::||||||      |||:::
Db        686 GTTACTAAGCCGTTCCGGTTTGAGGGAAACAACCGTGCAAAGCTTGCGGCACAAGGCCTC 745

Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
                  ||||||         ||||||||||:::::|||:::|||||||||||||||||
Db        746 GCTGAACTGAAGGATAGCGTCGATACGATGCTTGTGATCCCGAACCAAAACTTGTTCAAC 805

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
                :::::::::::   |||:::::::::::|||||||||   ||||||:::|||||||||
Db        806 ATGTCAAATGAGCGCACCTCGTTGATGGACGCATTCAGAATGGCGGACAATGTGCTTCTG 865

Qy        161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                :::|||||||||   ::::::|||||||:::|||   ||||||||||||||||||
Db        866 GACGGTGTCAAGAACATTTCGGATTTGATGGTGATGCCTGGGCTCATTAACCTT 919

RESULT 9
US-09-815-242-4715
; Sequence 4715, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4715
;   LENGTH: 1172
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-4715

Alignment Scores:
Pred. No.:              1.41e-53        Length:         1172
Score:                  480.00          Matches:        95
Percent Similarity:     73.14%          Conservative:   33
Best Local Similarity:  54.29%          Mismatches:     47
Query Match:            54.24%          Indels:         0
DB:                     10              Gaps:           0

US-09-770-509-2 (1-178) x US-09-815-242-4715 (1-1172)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              :::      ||||||||||||      ||||||||   ||||||||      |||
Db        103 ATGAATAATGTTGAATTTATCGCTATCAACACAGACGGTCAAGCTTTAAACTTATCTAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      ||||||   :::|||:::   :::|||:::||||||||||||||||:::
Db        163 GCTGAATCTAAAATCCAAATCGGTGAAAAATTAACACGTGGTTTAGGAGCAGGAGCTAAT 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||||:::::::|||||:::   :::      |||:::      :::|||
Db        223 CCTGAAATCGGTAAAAAAGCTGCAGAGGAATCTCGTGAACAAATTGAAGATGCAATCCAA 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                :::::|||:::|||:::|||   ||||||||||||||||   ||||||||||||
Db        283 GGTGCAGACATGGTATTTGTTACTTCTGGTATGGGTGGCGGAACTGGTACTGGTGCAGCA 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||:::|||   ||||||||||||||||||:::
Db        343 CCAGTCGTTGCTAAAATTGCAAAAGAAATGGGCGCATTAACTGTTGGTGTTGTAACTCGT 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||   |||||||||:::   :::
Db        403 CCATTTAGTTTTGAAGGACGTAAACGTCAAACTCAAGCTGCTGCTGGAGTAGAAGCTATG 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                  ||||||||||||||||||:::||||||      ||||||   :::   |||
Db        463 AAAGCTGCAGTAGATACATTAATCGTTATACCAAATGACCGTTTATTAGATATCGTTGAC 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||||||||   |||:::||||||||:::   ||||||:::||||||   :::|||||||
Db        523 AAATCTACGCCAATGATGGAAGCATTTAAAGAAGCTGACAACGTGTTACGCCAAGGTGTA 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::::|||||||||          |||   :::||||||
Db        583 CAAGGTATCTCAGACTTAATCGCTGTTTCTGGTGAAGTAAACTTA 627

RESULT 10
US-09-815-242-8625
; Sequence 8625, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8625
;   LENGTH: 1179
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1179)
```

```
US-09-815-242-8625

Alignment Scores:
Pred. No.:                1.42e-53        Length:          1179
Score:                    480.00          Matches:         95
Percent Similarity:       73.14%          Conservative:    33
Best Local Similarity:    54.29%          Mismatches:      47
Query Match:              54.24%          Indels:          0
DB:                       10              Gaps:            0

US-09-770-509-2 (1-178) x US-09-815-242-8625 (1-1179)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              :::      ||||||||||||      |||||||||  |||||||||      |||
Db        103 ATGAATAATGTTGAATTTATCGCTATCAACACAGACGGTCAAGCTTTAAACTTATCTAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      ||||||   :::|||:::   :::|||:::||||||||||||||||:::
Db        163 GCTGAATCTAAAATCCAAATCGGTGAAAAATTAACACGTGGTTTAGGAGCAGGAGCTAAT 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||::::::||||||:::   :::      |||:::      :::|||
Db        223 CCTGAAATCGGTAAAAAAGCTGCAGAGGAATCTCGTGAACAAATTGAAGATGCAATCCAA 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||:::|||   ||||||||||||||||||   ||||||||||||
Db        283 GGTGCAGACATGGTATTTGTTACTTCTGGTATGGGTGGCGGAACTGGTACTGGTGCAGCA 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||:::|||   ||||||||||||||||||:::
Db        343 CCAGTCGTTGCTAAAATTGCAAAAGAAATGGGCGCATTAACTGTTGGTGTTGTAACTCGT 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||   ||||||||:::   :::
Db        403 CCATTTAGTTTTGAAGGACGTAAACGTCAAACTCAAGCTGCTGCTGGAGTAGAAGCTATG 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              ||||||||||||||||||||:::||||||      ||||||   :::   |||
Db        463 AAAGCTGCAGTAGATACATTAATCGTTATACCAAATGACCGTTTATTAGATATCGTTGAC 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||||||||   |||:::||||||||:::   ||||||:::||||||   :::||||||
Db        523 AAATCTACGCCAATGATGGAAGCATTTAAAGAAGCTGACAACGTGTTACGCCAAGGTGTA 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::|||||||||         |||   :::||||||
Db        583 CAAGGTATCTCAGACTTAATCGCTGTTTCTGGTGAAGTAAACTTA 627

RESULT 11
US-08-781-986A-342
; Sequence 342, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
;   APPLICANT:  Charles Kunsch
;   TITLE OF INVENTION:  Staphylococcus aureus Polynucleotides and Sequences
;   NUMBER OF SEQUENCES:  5255
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Human Genome Sciences, Inc.
;     STREET:  9410 Key West Avenue
;     CITY:  Rockville
;     STATE:  Maryland
;     COUNTRY:  USA
;     ZIP:  20850
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;     COMPUTER:  HP Vectra 486/33
;     OPERATING SYSTEM:  MSDOS version 6.2
;     SOFTWARE:  ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/781,986A
;     FILING DATE:
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Benson, Bob
;     REGISTRATION NUMBER:  30,446
;     REFERENCE/DOCKET NUMBER:  PB248PP
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (301) 309-8504
;     TELEFAX:  (301) 309-8512
; INFORMATION FOR SEQ ID NO:  342:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  3931 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  double
;     TOPOLOGY:  linear
US-08-781-986A-342

Alignment Scores:
Pred. No.:                8.05e-53        Length:          3931
Score:                    480.00          Matches:         95
Percent Similarity:       73.14%          Conservative:    33
Best Local Similarity:    54.29%          Mismatches:      47
Query Match:              54.24%          Indels:          0
DB:                       7               Gaps:            0

US-09-770-509-2 (1-178) x US-08-781-986A-342 (1-3931)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              :::      ||||||||||||      |||||||||  |||||||||      |||
Db        684 ATGAATAATGTTGAATTTATCGCTATCAACACAGACGGTCAAGCTTTAAACTTATCTAAA 743

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      ||||||   :::|||:::   :::|||:::||||||||||||||||:::
Db        744 GCTGAATCTAAAATCCAAATCGGTGAAAAATTAACACGTGGTTTAGGAGCAGGAGCTAAT 803

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||::::::||||||:::   :::      |||:::      :::|||
Db        804 CCTGAAATCGGTAAAAAAGCTGCAGAGGAATCTCGTGAACAAATTGAAGATGCAATCCAA 863

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||:::|||   ||||||||||||||||||   ||||||||||||
Db        864 GGTGCAGACATGGTATTTGTTACTTCTGGTATGGGTGGCGGAACTGGTACTGGTGCAGCA 923

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||:::|||   ||||||||||||||||||:::
Db        924 CCAGTCGTTGCTAAAATTGCAAAAGAAATGGGCGCATTAACTGTTGGTGTTGTAACTCGT 983

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||   ||||||||:::   :::
Db        984 CCATTTAGTTTTGAAGGACGTAAACGTCAAACTCAAGCTGCTGCTGGAGTAGAAGCTATG 1043

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              ||||||||||||||||||||:::||||||      ||||||   :::   |||
Db       1044 AAAGCTGCAGTAGATACATTAATCGTTATACCAAATGACCGTTTATTAGATATCGTTGAC 1103

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||||||||   |||:::||||||||:::   ||||||:::||||||   :::||||||
Db       1104 AAATCTACGCCAATGATGGAAGCATTTAAAGAAGCTGACAACGTGTTACGCCAAGGTGTA 1163

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::|||||||||         |||   :::||||||
Db       1164 CAAGGTATCTCAGACTTAATCGCTGTTTCTGGTGAAGTAAACTTA 1208

RESULT 12
US-09-754-608-1
; Sequence 1, Application US/09754608
; Patent No. US20020004580A1
; GENERAL INFORMATION:
;  APPLICANT: Fueyo, Joanna Lynn
;  APPLICANT:  Lonetto, Michael A.
;  APPLICANT:  Pearce, Kenneth
;  TITLE OF INVENTION: ftsZ
;  FILE REFERENCE: GM10068
```

```
;  CURRENT APPLICATION NUMBER: US/09/754,608
;  CURRENT FILING DATE:  2001-01-04
;  PRIOR APPLICATION NUMBER: 09/120,426
;  PRIOR APPLICATION NUMBER: 1998-07-22
;  PRIOR APPLICATION NUMBER: 60/055,720
;  PRIOR FILING DATE: 1997-08-12
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 1260
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
US-09-754-608-1

Alignment Scores:
Pred. No.:                1.77e-52      Length:          1260
Score:                    472.00        Matches:         91
Percent Similarity:       72.83%        Conservative:    35
Best Local Similarity:    52.60%        Mismatches:      47
Query Match:              53.33%        Indels:          0
DB:                       10            Gaps:            0

US-09-770-509-2 (1-178) x US-09-754-608-1 (1-1260)

Qy          6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
              ||||||||||||||||   ||||||||||||   |||||||||      :::   |||
Db        112 GGCGTAGAATTTATCGCAGCAAACACAGATGTACAAGCATTGAGTAGTACAAAAGCTGAG 171

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
                 |||   ||||||      :::|||:::|||||||||||||||   :::|||||||
Db        172 ACTGTTATTCAGTTGGGACCTAAATTGACTCGTGGTTTGGGTGCAGGAGGTCAACCTGAG 231

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
              :::|||::::::::::||||||:::   :::      :::          :::      :::
Db        232 GTTGGTCGTAAAGCCGCTGAAGAAAGCGAAGAAACACTGACGGAAGCTATTAGTGGTGCC 291

Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
              :::|||:::||||||||||   ||||||||||||||||:::   ||||||||||||||||||
Db        292 GATATGGTCTTCATCACTGCTGGTATGGGAGGAGGCTCTGGAACTGGAGCTGCTCCTGTT 351

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
              :::|||   :::|||::::::||||||   ||||||||||||||||||:::   ||||||
Db        352 ATTGCTCGTATCGCCAAAGATTTAGGTGCGCTTACAGTTGGTGTTGTAACACGTCCCTTT 411

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
                |||||      |||   :::   |||      |||:::   :::|||   :::
Db        412 GGTTTTGAAGGAAGTAAGCGTGGACAATTTGCTGTAGAAGGAATCAATCAACTTCGTGAG 471

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
              :::||||||||||||:::::::::   |||   |||||||||   :::   ||||||
Db        472 CATGTAGACACTCTATTGATTATCTCAAACAACAATTTGCTTGAAATTGTTGATAAGAAA 531

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
              |||   :::||||||||||       ||||||:::|||||||   :::|||||:::|||
Db        532 ACACCGCTTTTGGAGGCTCTTAGCGAAGCGGATAACGTTCTTCGTCAAGGTGTTCAAGGG 591

Qy        166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::||||||||||||      ||||||||||||||||||
Db        592 ATTACCGATTTGATTACCAATCCAGGATTGATTAACCTT 630

RESULT 13
US-09-815-242-9381
; Sequence 9381, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9381
;   LENGTH: 1260
;   TYPE: DNA
;   ORGANISM: Streptococcus pneumoniae
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1260)
US-09-815-242-9381

Alignment Scores:
Pred. No.:                1.77e-52      Length:          1260
Score:                    472.00        Matches:         91
Percent Similarity:       72.83%        Conservative:    35
Best Local Similarity:    52.60%        Mismatches:      47
Query Match:              53.33%        Indels:          0
DB:                       10            Gaps:            0

US-09-770-509-2 (1-178) x US-09-815-242-9381 (1-1260)

Qy          6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
              ||||||||||||||||   ||||||||||||   |||||||||      :::   |||
Db        112 GGCGTAGAATTTATCGCAGCAAACACAGATGTACAAGCATTGAGTAGTACAAAAGCTGAG 171

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
                 |||   ||||||      :::|||:::|||||||||||||||   :::|||||||
Db        172 ACTGTTATTCAGTTGGGACCTAAATTGACTCGTGGTTTGGGTGCAGGAGGTCAACCTGAG 231

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
              :::|||::::::::::||||||:::   :::      :::          :::      :::
Db        232 GTTGGTCGTAAAGCCGCTGAAGAAAGCGAAGAAACACTGACGGAAGCTATTAGTGGTGCC 291

Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
              :::|||:::||||||||||   ||||||||||||||||:::   ||||||||||||||||||
Db        292 GATATGGTCTTCATCACTGCTGGTATGGGAGGAGGCTCTGGAACTGGAGCTGCTCCTGTT 351

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
              :::|||   :::|||::::::||||||   ||||||||||||||||||:::   ||||||
Db        352 ATTGCTCGTATCGCCAAAGATTTAGGTGCGCTTACAGTTGGTGTTGTAACACGTCCCTTT 411

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
                |||||      |||   :::   |||      |||:::   :::|||   :::
Db        412 GGTTTTGAAGGAAGTAAGCGTGGACAATTTGCTGTAGAAGGAATCAATCAACTTCGTGAG 471

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
              :::||||||||||||:::::::::   |||   |||||||||   :::   ||||||
Db        472 CATGTAGACACTCTATTGATTATCTCAAACAACAATTTGCTTGAAATTGTTGATAAGAAA 531

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
              |||   :::||||||||||       ||||||:::|||||||   :::|||||:::|||
Db        532 ACACCGCTTTTGGAGGCTCTTAGCGAAGCGGATAACGTTCTTCGTCAAGGTGTTCAAGGG 591
```

```
Qy        166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||||||||||||      ||||||||||||||||||
Db        592 ATTACCGATTTGATTACCAATCCAGGATTGATTAACCTT 630
```

RESULT 14
US-09-738-626-2360
; Sequence 2360, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 2360
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-2360

Alignment Scores:

Pred. No.:	3.49e-52	Length:	1326
Score:	470.00	Matches:	92
Percent Similarity:	73.14%	Conservative:	36
Best Local Similarity:	52.57%	Mismatches:	47
Query Match:	53.11%	Indels:	0
DB:	9	Gaps:	0

US-09-770-509-2 (1-178) x US-09-738-626-2360 (1-1326)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||||||||||      |||||||||   |||||||||      |||
Db         94 CTCAAAGGCGTGGAGTTCATCGCGGTGAACACCGACTCGCAGGCTCTCATGTTCTCTGAT 153

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||::::::   |||:::||||||||||||||||:::
Db        154 GCCGACGTAAAGCTCGATATCGGACGTGAAGCTACCCGTGGTCTTGGTGCCGGCGCGAAC 213

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::   |||||||||      |||   :::|||:::      ::::::
Db        214 CCAGAAGTTGGACGTGCCTCGGCAGAGGATCACAAGAACGAAATCGAAGAAACCATCAAG 273

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||:::|||   |||   ||||||||||||   ||||||||||||
Db        274 GGCGCCGACATGGTCTTCGTTACCGCCGGCGAAGGTGGTGGCACCGGAACTGGTGCTGCA 333

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::::::::|||   |||||||:::||||||||:::
Db        334 CCAGTCGTGGCAGGGATCGCCAAGAAGATGGGCGCACTGACCATTGGTGTTGTGACCAAG 393

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||||||||   |||      |||:::      |||
Db        394 CCTTTCGAGTTCGAAGGCCGTCGCCGTACTCGCCAGGCAGAAGAAGGCATCGCAGCACTG 453

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::      ||||||||||||||||:::||||||      ||||||   |||   |||

```

```
Db        454 AAGGAGGTCTGCGACACCCTCATCGTTATTCCAAACGACCGCCTGCTTGAGCTGGGCGAT 513

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
                 :::   :::::::::|||||||||||||   ||||||:::||||||      ||||||
Db        514 GCGAACCTGTCCATCATGGAAGCGTTCCGCGCAGCCGATGAAGTTCTCCACAATGGTGTT 573

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::|||:::||||||      ||||||:::||||||:::
Db        574 CAGGGTATTACCAACCTGATCACCATCCCTGGTGTGATCAACGTG 618
```

RESULT 15
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Alignment Scores:

Pred. No.:	2.74e-47	Length:	3309400
Score:	470.00	Matches:	92
Percent Similarity:	73.14%	Conservative:	36
Best Local Similarity:	52.57%	Mismatches:	47
Query Match:	53.11%	Indels:	0
DB:	9	Gaps:	0

US-09-770-509-2 (1-178) x US-09-738-626-1 (1-3309400)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||||||||||      |||||||||   |||||||||      |||
Db    2280122 CTCAAAGGCGTGGAGTTCATCGCGGTGAACACCGACTCGCAGGCTCTCATGTTCTCTGAT 2280063

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||::::::   |||:::||||||||||||||||:::
Db    2280062 GCCGACGTAAAGCTCGATATCGGACGTGAAGCTACCCGTGGTCTTGGTGCCGGCGCGAAC 2280003

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::   |||||||||      |||   :::|||:::      ::::::
Db    2280002 CCAGAAGTTGGACGTGCCTCGGCAGAGGATCACAAGAACGAAATCGAAGAAACCATCAAG 2279943

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||:::|||   |||   ||||||||||||   ||||||||||||
Db    2279942 GGCGCCGACATGGTCTTCGTTACCGCCGGCGAAGGTGGTGGCACCGGAACTGGTGCTGCA 2279883

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::::::::|||   |||||||:::||||||||:::
Db    2279882 CCAGTCGTGGCAGGGATCGCCAAGAAGATGGGCGCACTGACCATTGGTGTTGTGACCAAG 2279823
```

```
Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||||||||   |||      |||:::      |||
Db    2279822 CCTTTCGAGTTCGAAGGCCGTCGCCGTACTCGCCAGGCAGAAGAAGGCATCGCAGCACTG 2279763

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::      ||||||||||||||||:::||||||      ||||||   |||   |||
Db    2279762 AAGGAGGTCTGCGACACCCTCATCGTTATTCCAAACGACCGCCTGCTTGAGCTGGGCGAT 2279703

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
                 :::   :::::::::||||||||||||   ||||||:::|||||      ||||||
Db    2279702 GCGAACCTGTCCATCATGGAAGCGTTCCGCGCAGCCGATGAAGTTCTCCACAATGGTGTT 2279643

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::|||:::||||||      ||||||:::||||||:::
Db    2279642 CAGGGTATTACCAACCTGATCACCATCCCTGGTGTGATCAACGTG 2279598
```

Search completed: June 2, 2003, 13:17:12
Job time : 7750 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:52:57 ; Search time 2738 Seconds
(without alignments)
1634.537 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09770509/runat_27052003_083651_18863/app_query.fasta_1.327
-DB=Pending_Patents_NA_Main -QFMT=fastap -SUFFIX=rnpm -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09770509_@CGN_1_1_2950_@runat_27052003_083651_18863 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		885	100.0	535	30	US-09-770-509-1	Sequence 1, Appli
2		550	62.1	1473	38	US-10-015-127-5607	Sequence 5607, Ap
3	c	550	62.1	24740	38	US-10-015-127-522	Sequence 522, App
4		509	57.5	950	36	US-09-974-300-4664	Sequence 4664, Ap
5	c	509	57.5	2074	71	US-60-279-526-159	Sequence 159, App
6		506	57.2	111309	1	PCT-US98-12764-3	Sequence 3, Appli

	7	506	57.2	111309	32	US-09-830-228-3 Sequence 3, Appli
	8	506	57.2	910715	1	PCT-US98-12764-1 Sequence 1, Appli
	9	506	57.2	910715	32	US-09-830-228-1 Sequence 1, Appli
c	10	505	57.1	7375	26	US-09-663-779-1189 Sequence 1189, Ap
	11	503	56.8	1864	29	US-09-754-468-42 Sequence 42, Appl
	12	502	56.7	1134	36	US-09-974-300-180 Sequence 180, App
	13	496	56.0	3993	48	US-60-045-649-854 Sequence 854, App
	14	496	56.0	3993	48	US-60-046-653-853 Sequence 853, App
	15	495	55.9	1233	1	PCT-US02-03987-6542 Sequence 6542, Ap
	16	495	55.9	1233	31	US-09-815-242-6542 Sequence 6542, Ap
	17	495	55.9	1233	39	US-10-072-851-6542 Sequence 6542, Ap
	18	495	55.9	1239	1	PCT-US02-03987-3854 Sequence 3854, Ap
	19	495	55.9	1239	31	US-09-815-242-3854 Sequence 3854, Ap
	20	495	55.9	1239	39	US-10-072-851-3854 Sequence 3854, Ap
	21	495	55.9	1308	15	US-09-134-000-1253 Sequence 1253, Ap
	22	495	55.9	3904	50	US-60-068-217-757 Sequence 757, App
c	23	495	55.9	32768	14	US-09-070-927-128 Sequence 128, App
c	24	495	55.9	32768	14	US-09-070-927A-128 Sequence 128, App
	25	494	55.8	1245	15	US-09-107-532-3641 Sequence 3641, Ap
	26	494	55.8	1245	15	US-09-107-532A-3641 Sequence 3641, Ap
	27	494	55.8	13986	49	US-60-050-444-716 Sequence 716, App
	28	494	55.8	13986	50	US-60-068-186-720 Sequence 720, App
	29	485	54.8	1140	28	US-09-712-363-75 Sequence 75, Appl
c	30	485	54.8	37586	29	US-09-754-468-27 Sequence 27, Appl
c	31	482	54.5	1662	19	US-09-514-000-5249 Sequence 5249, Ap
	32	482	54.5	1752	29	US-09-739-449-3270 Sequence 3270, Ap
	33	482	54.5	1752	31	US-09-803-110-3270 Sequence 3270, Ap
c	34	482	54.5	1867	60	US-60-168-139-1162 Sequence 1162, Ap
c	35	482	54.5	27239	19	US-09-514-000-330 Sequence 330, App
c	36	482	54.5	184668	29	US-09-739-449-207 Sequence 207, App
c	37	482	54.5	184668	31	US-09-803-110-207 Sequence 207, App
	38	481	54.4	1423	30	US-09-770-509-9 Sequence 9, Appli
	39	480	54.2	1158	14	US-09-037-934-12 Sequence 12, Appl
	40	480	54.2	1169	24	US-09-620-608-1462 Sequence 1462, Ap
	41	480	54.2	1169	58	US-60-144-883-1462 Sequence 1462, Ap
	42	480	54.2	1170	16	US-09-266-541-1 Sequence 1, Appli
	43	480	54.2	1172	1	PCT-US02-03987-4715 Sequence 4715, Ap
	44	480	54.2	1172	31	US-09-815-242-4715 Sequence 4715, Ap
	45	480	54.2	1172	39	US-10-072-851-4715 Sequence 4715, Ap

ALIGNMENTS

```
RESULT 1
US-09-770-509-1
; Sequence 1, Application US/09770509
; GENERAL INFORMATION:
;  APPLICANT: Katagiri, F.
;  TITLE OF INVENTION: OOMYCETE FTSZ-MT AS A TARGET FOR
;  TITLE OF INVENTION:  ANTIMICROBIAL-SPECIFIC BIOCIDES
;  FILE REFERENCE: NADII.018A
;  CURRENT APPLICATION NUMBER: US/09/770,509
;  CURRENT FILING DATE:  2001-01-25
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 535
;   TYPE: DNA
;   ORGANISM: Phytophthora infestans
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)...(535)
US-09-770-509-1

Alignment Scores:
Pred. No.:              3.97e-84        Length:         535
Score:                  885.00          Matches:        178
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     30              Gaps:           0

US-09-770-509-2 (1-178) x US-09-770-509-1 (1-535)
```

```
Qy          1 AlaSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GCGTCGCAATTGGAAGGTGTGGAGTTCATTGTAGCCAACACAGACTGTCAGGCTCTGGGA 61

Qy         21 ArgSerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CGCTCGCTGGCGCCGCACAAGATCACGCTGGGCAAAGATATCACCAAGGGACTAGGAGCT 121

Qy         41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GGATCCAAACCTGAGCTGGGTAAACGCTCTGCGGAACAGCAGAAAGTGGATATCCAACGG 181

Qy         61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 ATGTTACAGGACAGCAACATGCTGTTTATCACGGGCGGAATGGGCGGCGGAACCTGCACA 241

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GGAGCCGCACCTGTCGTGGCCAGTGTAGCCAGGGAGCTGGGGATCCTAACGGTCGGAGTA 301

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GTAAGCACACCGTTCCGATCCGAAGGACCCAATCGCACTCGTCTGGCCAATGCTGGAGTA 361

Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AAAGAACTGGCCAAGTACGTCGACACCTTAATTGTCGTGCCCAACCAGAACTTGCTGGCT 421

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTGGCAGACAAGAGCACGACCATGTTGGAAGCCTTCCGGTATGCCGACGACGTGCTGCTT 481

Qy        161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAAGGAGTTAAAGGTGTCACGGACTTGATCGTTCGCCCGGGACTTATCAATTTG 535
```

```
RESULT 2
US-10-015-127-5607
; Sequence 5607, Application US/10015127
; GENERAL INFORMATION:
;  APPLICANT: Bower, Stanley G.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  TITLE OF INVENTION: Sphingomonas elodea genome sequences and uses thereof
;  FILE REFERENCE: 38-10(15806)B
;  CURRENT APPLICATION NUMBER: US/10/015,127
;  CURRENT FILING DATE:  2001-10-29
;  PRIOR APPLICATION NUMBER: US 60/252,455
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 14357
; SEQ ID NO 5607
;   LENGTH: 1473
;   TYPE: DNA
;   ORGANISM: Sphingomonas elodea
US-10-015-127-5607

Alignment Scores:
Pred. No.:              5.61e-48        Length:         1473
Score:                  550.00          Matches:        105
Percent Similarity:     79.10%          Conservative:   35
Best Local Similarity:  59.32%          Mismatches:     37
Query Match:            62.15%          Indels:         0
DB:                     38              Gaps:           0

US-09-770-509-2 (1-178) x US-10-015-127-5607 (1-1473)
```

```
Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              :::::::::::::||||||||||||||:::|||||||||||||||||   |||||||||   :::
Db        106 GCGGAGGTGCAGGGGGTCGAGTTCCTGGTCGCCAACACCGATGCGCAGGCGCTGAAGCAG 165

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||:::||||||   :::|||   ||||||      |||||||:::|||||||||||||||
```

```
Db        166 TCGGTTGCGCCGCAGCGCATTCAGCTGGGCGCGAAGATCACGCAGGGCCTGGGCGCGGGC 225

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              |||:::|||||:::|||:::   :::|||||:::          :::|||:::
Db        226 TCGCGGCCCGAAATCGGCCGTGCGGCGGCCGAGGAGACGATCGAGCAGGTCCAGCAGTCG 285

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::   :::::|||   ||||||      ||||||||||||||||   ||||||
Db        286 CTCGAAGGCGCGCACATGTGCTTCATCGCCGCCGGCATGGGCGGCGGCACCGGCACGGGT 345

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||:::|||      ||||||::::::||||||||||||||||||||||||
Db        346 GCGGCACCGGTGATCGCCAAGGCGGCGCGCGACATGGGCATCCTGACGGTGGGCGTCGTC 405

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||      |||   :::   |||   ||||||::::::
Db        406 ACCAAGCCGTTCGCCTTCGAAGGCAAGCGCCGCGCGCAATCGGCCGAGGCGGGCATCGAG 465

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              ||||||   ||||||||||||||||||||:::|||||||||||||||      :::
Db        466 GAACTGCAGAAGTACGTCGACACCCTGATCGTCATCCCCAACCAGAATCTCTTCCTGATC 525

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              |||:::   :::||||||      ||||||||      ||||||:::||||||   :::
Db        526 GCCAATGCGAACACGACCTTCAAGGAAGCGTTCGAGATGGCCGACGAGGTGCTGCAGCAG 585

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||:::|||||||||:::|||   ||||||||||||||||
Db        586 GGTGTGCGCGGCATCACCGACCTGATGGTGATGCCCGGCCTGATCAACCTC 636

RESULT 3
US-10-015-127-522/c
; Sequence 522, Application US/10015127
; GENERAL INFORMATION:
;  APPLICANT: Bower, Stanley G.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  TITLE OF INVENTION: Sphingomonas elodea genome sequences and uses thereof
;  FILE REFERENCE: 38-10(15806)B
;  CURRENT APPLICATION NUMBER: US/10/015,127
;  CURRENT FILING DATE:  2001-10-29
;  PRIOR APPLICATION NUMBER: US 60/252,455
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 14357
; SEQ ID NO 522
;   LENGTH: 24740
;   TYPE: DNA
;   ORGANISM: Sphingomonas elodea
US-10-015-127-522

Alignment Scores:
Pred. No.:              1.68e-46        Length:         24740
Score:                  550.00          Matches:        105
Percent Similarity:     79.10%          Conservative:   35
Best Local Similarity:  59.32%          Mismatches:     37
Query Match:            62.15%          Indels:         0
DB:                     38              Gaps:           0

US-09-770-509-2 (1-178) x US-10-015-127-522 (1-24740)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              ::::::::::::|||||||||||:::|||||||||||||||   |||||||||   :::
Db       3486 GCGGAGGTGCAGGGGGTCGAGTTCCTGGTCGCCAACACCGATGCGCAGGCGCTGAAGCAG 3427

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||:::||||||   :::|||   ||||||      ||||||:::|||||||||||||||
Db       3426 TCGGTTGCGCCGCAGCGCATTCAGCTGGGCGCGAAGATCACGCAGGGCCTGGGCGCGGGC 3367

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              |||:::|||||:::|||:::   :::|||||:::          :::|||:::
Db       3366 TCGCGGCCCGAAATCGGCCGTGCGGCGGCCGAGGAGACGATCGAGCAGGTCCAGCAGTCG 3307

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::   :::::|||   ||||||      ||||||||||||||||   ||||||
Db       3306 CTCGAAGGCGCGCACATGTGCTTCATCGCCGCCGGCATGGGCGGCGGCACCGGCACGGGT 3247

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||:::|||      ||||||::::::||||||||||||||||||||||||
Db       3246 GCGGCACCGGTGATCGCCAAGGCGGCGCGCGACATGGGCATCCTGACGGTGGGCGTCGTC 3187

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||      |||   :::   |||   ||||||::::::
Db       3186 ACCAAGCCGTTCGCCTTCGAAGGCAAGCGCCGCGCGCAATCGGCCGAGGCGGGCATCGAG 3127

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              ||||||   ||||||||||||||||||||:::|||||||||||||||      :::
Db       3126 GAACTGCAGAAGTACGTCGACACCCTGATCGTCATCCCCAACCAGAATCTCTTCCTGATC 3067

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              |||:::   :::||||||      ||||||||      ||||||:::||||||   :::
Db       3066 GCCAATGCGAACACGACCTTCAAGGAAGCGTTCGAGATGGCCGACGAGGTGCTGCAGCAG 3007

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||:::|||||||||:::|||   ||||||||||||||||
Db       3006 GGTGTGCGCGGCATCACCGACCTGATGGTGATGCCCGGCCTGATCAACCTC 2956

RESULT 4
US-09-974-300-4664
; Sequence 4664, Application US/09974300
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4664
;   LENGTH: 950
;   TYPE: DNA
;   ORGANISM: Bacillus clausii
US-09-974-300-4664

Alignment Scores:
Pred. No.:              7.56e-44        Length:         950
Score:                  509.00          Matches:        101
Percent Similarity:     74.86%          Conservative:   30
Best Local Similarity:  57.71%          Mismatches:     44
Query Match:            57.51%          Indels:         0
DB:                     36              Gaps:           0

US-09-770-509-2 (1-178) x US-09-974-300-4664 (1-950)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||:::|||||      |||||||||   |||||||||      |||
Db        103 CTGCAAGGTGTTGATTTTATTGCAGTCAATACAGATGCACAAGCACTCCATTTATCAAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   ||||||      :::|||:::|||||||||||||:::
Db        163 GCGGAAAAGAAGTTGCAACTTGGCGGCAAACTTACTCGCGGCCTTGGCGCTGGCGCCAAC 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||||:::::::|||||:::   :::      ::::::   :::|||
Db        223 CCAGAAATAGGCAAGAAAGCGGCAGAGGAAAGCAGAGAACAGCTGGAAGAAGTTCTTACA 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||:::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        283 GGTTCTGATATGGTCTTTATTACTGCCGGAATGGGAGGAGGAACAGGCACCGGAGCAGCT 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
```

```
Db        343 CCTGTCATTGCTGAAGTGGCAAAAGAGCTTGGCGCATTGACGGTCGGTGTTGTGACACGC 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
Db        403 CCTTTCTCTTTTGAAGGCCGTAAGCGGCAAAACCAAGCGATCTCTGGGATTGCAGCGTTA 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
Db        463 AAAGAAAAAGTGGACACATTGATCGTGATTCCAAATGACCGCTTATTGGAAATGGTTGAT 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
Db        523 AAAAATACGCCGATGCTTGAAGCGTTCCGCGAGGCGGATAATGTATTGCGCCAAGGAGTT 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
Db        583 CAAGGCATTTCCGATTTAATTGCCACGCCTGGGTTGATCAACCTT 627

RESULT 5
US-60-279-526-159/c
; Sequence 159, Application US/60279526
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  APPLICANT:  Sorokine, Alexei
;  APPLICANT:  Bolotine, Alexandre
;  APPLICANT:  Lapidus, Alla
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 5963.000-US
;  CURRENT APPLICATION NUMBER: US/60/279,526
;  CURRENT FILING DATE:  2001-03-27
;  NUMBER OF SEQ ID NOS: 2232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159
;   LENGTH: 2074
;   TYPE: DNA
;   ORGANISM: B. clausii
US-60-279-526-159

Alignment Scores:
Pred. No.:              1.94e-43      Length:          2074
Score:                  509.00        Matches:         101
Percent Similarity:     74.86%        Conservative:    30
Best Local Similarity:  57.71%        Mismatches:      44
Query Match:            57.51%        Indels:          0
DB:                     71            Gaps:            0

US-09-770-509-2 (1-178) x US-60-279-526-159 (1-2074)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
Db        857 CTGCAAGGTGTTGATTTTATTGCAGTCAATACAGATGCACAAGCACTCCATTTATCAAAA 798

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
Db        797 GCGGAAAAGAAGTTGCAACTTGGCGGCAAACTTACTCGCGGCCTTGGCGCTGGCGCCAAC 738

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
Db        737 CCAGAAATAGGCAAGAAAGCGGCAGAGGAAAGCAGAGAACAGCTGGAAGAAGTTCTTACA 678

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
Db        677 GGTTCTGATATGGTCTTTATTACTGCCGGAATGGGAGGAGGAACAGGCACCGGAGCAGCT 618

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
Db        617 CCTGTCATTGCTGAAGTGGCAAAAGAGCTTGGCGCATTGACGGTCGGTGTTGTGACACGC 558

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
Db        557 CCTTTCTCTTTTGAAGGCCGTAAGCGGCAAAACCAAGCGATCTCTGGGATTGCAGCGTTA 498

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
Db        497 AAAGAAAAAGTGGACACATTGATCGTGATTCCAAATGACCGCTTATTGGAAATGGTTGAT 438

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
Db        437 AAAAATACGCCGATGCTTGAAGCGTTCCGCGAGGCGGATAATGTATTGCGCCAAGGAGTT 378

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
Db        377 CAAGGCATTTCCGATTTAATTGCCACGCCTGGGTTGATCAACCTT 333

RESULT 6
PCT-US98-12764-3
; Sequence 3, Application PC/TUS9812764
;  GENERAL INFORMATION:
;    APPLICANT:  Gil Choi et. al.
;    TITLE OF INVENTION:  Borrelia burgdorferi Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  155
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US98/12764
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB370PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  111309 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
PCT-US98-12764-3

Alignment Scores:
Pred. No.:              4.92e-41      Length:          111309
Score:                  506.00        Matches:         101
Percent Similarity:     74.42%        Conservative:    27
Best Local Similarity:  58.72%        Mismatches:      44
Query Match:            57.18%        Indels:          0
DB:                     1             Gaps:            0

US-09-770-509-2 (1-178) x PCT-US98-12764-3 (1-111309)

Qy          7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
Db      12947 GTTGAATTTATTGTGGCTAATACCGATCTTCAGGCTCTCCAAACTTCTATTGCTCCCATA 13006

Qy         27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
Db      13007 AAAATTGCCCTTGGAGCAAAAGTTACAGCAGGGCTTGGTGCTGGGGGAAAGCCTGAGATT 13066
```

```
Qy         47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
              |||:::   :::|||||||:::            |||:::      |||      ::::::
Db      13067 GGACAAGCTGCAGCAGAGGAAGACATAGATGTTATACGAAATCATCTTTCTGGTGCCGAT 13126

Qy         67 MetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValVal 86
              |||:::|||||||||   ||||||||||||||||||   ||||||||||||||||||:::
Db      13127 ATGGTGTTTATTACTGCTGGTATGGGGGGCGGGACAGGAACCGGAGCAGCTCCAGTTATT 13186

Qy         87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArg 106
              |||   ||||||:::|||||||||||||||||||||||||||||:::   ||||||:::
Db      13187 GCGCAAGTTGCAAAAGAGCTTGGTATTTTAACAGTTGGAGTTGTAACAAAGCCTTTTAAG 13246

Qy        107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyr 126
                 |||||||||   :::   |||||||||      |||:::      |||   |||
Db      13247 TTTGAAGGTCCTAAGAAGTTGAGACTTGCTGAGCAGGGAATAAATAACTTAAGAAAGTCT 13306

Qy        127 ValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThr 146
              ||||||||||||||:::::|||||||||   ||||||   :::   ||||||   |||
Db      13307 GTAGATACATTGATCATTATTCCAAATCAAAAGCTTTTAACTGTTGTTGACAAAAGAACC 13366

Qy        147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
              |||:::   :::|||||||:::   |||||||||||||||      ||||||:::|||:::
Db      13367 ACCATTAAAGATGCTTTTAAGCGTGCAGATGATGTTCTTAGAATGGGCGTTCAAGGTATT 13426

Qy        167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                 ||||||:::      |||   :::|||:::
Db      13427 GCAGGGCTTATTATTGAGCATGGAGAGGTTAATATT 13462

RESULT 7
US-09-830-228-3
; Sequence 3, Application US/09830228
;   GENERAL INFORMATION:
;        APPLICANT: Gil Choi et. al.
;        TITLE OF INVENTION: Borrelia burgdorferi Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 155
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/830,228
;             FILING DATE: 24-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/US98/12764
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brookes, A. Anders
;             REGISTRATION NUMBER: 36,373
;             REFERENCE/DOCKET NUMBER: PB370PCT
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 111309 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-830-228-3

Alignment Scores:
Pred. No.:              4.92e-41        Length:         111309
Score:                  506.00          Matches:        101
Percent Similarity:     74.42%          Conservative:   27
Best Local Similarity:  58.72%          Mismatches:     44
Query Match:            57.18%          Indels:         0
DB:                     32              Gaps:           0

US-09-770-509-2 (1-178) x US-09-830-228-3 (1-111309)

Qy          7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
              |||||||||||||||||||||||||   ||||||||      |||:::||||||
Db      12947 GTTGAATTTATTGTGGCTAATACCGATCTTCAGGCTCTCCAAACTTCTATTGCTCCCATA 13006

Qy         27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
              ||||||   ||||||      :::|||   ||||||||||||||   |||||||||:::
Db      13007 AAAATTGCCCTTGGAGCAAAAGTTACAGCAGGGCTTGGTGCTGGGGGAAAGCCTGAGATT 13066

Qy         47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
              |||:::   :::|||||||:::            |||:::      |||      ::::::
Db      13067 GGACAAGCTGCAGCAGAGGAAGACATAGATGTTATACGAAATCATCTTTCTGGTGCCGAT 13126

Qy         67 MetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValVal 86
              |||:::|||||||||   ||||||||||||||||||   ||||||||||||||||||:::
Db      13127 ATGGTGTTTATTACTGCTGGTATGGGGGGCGGGACAGGAACCGGAGCAGCTCCAGTTATT 13186

Qy         87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArg 106
              |||   ||||||:::|||||||||||||||||||||||||||||:::   ||||||:::
Db      13187 GCGCAAGTTGCAAAAGAGCTTGGTATTTTAACAGTTGGAGTTGTAACAAAGCCTTTTAAG 13246

Qy        107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyr 126
                 |||||||||   :::   |||||||||      |||:::      |||   |||
Db      13247 TTTGAAGGTCCTAAGAAGTTGAGACTTGCTGAGCAGGGAATAAATAACTTAAGAAAGTCT 13306

Qy        127 ValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThr 146
              ||||||||||||||:::::|||||||||   ||||||   :::   ||||||   |||
Db      13307 GTAGATACATTGATCATTATTCCAAATCAAAAGCTTTTAACTGTTGTTGACAAAAGAACC 13366

Qy        147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
              |||:::   :::|||||||:::   |||||||||||||||      ||||||:::|||:::
Db      13367 ACCATTAAAGATGCTTTTAAGCGTGCAGATGATGTTCTTAGAATGGGCGTTCAAGGTATT 13426

Qy        167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                 ||||||:::      |||   :::|||:::
Db      13427 GCAGGGCTTATTATTGAGCATGGAGAGGTTAATATT 13462

RESULT 8
PCT-US98-12764-1
; Sequence 1, Application PC/TUS9812764
;  GENERAL INFORMATION:
;    APPLICANT:  Gil Choi et. al.
;    TITLE OF INVENTION:  Borrelia burgdorferi Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  155
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US98/12764
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB370PCT
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  910715 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
PCT-US98-12764-1

Alignment Scores:
Pred. No.:              6.2e-40        Length:         910715
Score:                  506.00         Matches:        101
Percent Similarity:     74.42%         Conservative:   27
Best Local Similarity:  58.72%         Mismatches:     44
Query Match:            57.18%         Indels:         0
DB:                     1              Gaps:           0

US-09-770-509-2 (1-178) x PCT-US98-12764-1 (1-910715)

Qy          7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
              ||||||||||||||||||||||||||||  |||||||||      |||:::||||||
Db     602333 GTTGAATTTATTGTGGCTAATACCGATCTTCAGGCTCTCCAAACTTCTATTGCTCCCATA 602392

Qy         27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
              ||||||  ||||||      :::|||  ||||||||||||||  |||||||||:::
Db     602393 AAAATTGCCCTTGGAGCAAAAGTTACAGCAGGGCTTGGTGCTGGGGGAAAGCCTGAGATT 602452

Qy         47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
              |||:::  :::|||||:::            |||:::      |||      ::::::
Db     602453 GGACAAGCTGCAGCAGAGGAAGACATAGATGTTATACGAAATCATCTTTCTGGTGCCGAT 602512

Qy         67 MetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValVal 86
              |||:::|||||||||  ||||||||||||||||||  ||||||||||||||||||:::
Db     602513 ATGGTGTTTATTACTGCTGGTATGGGGGGCGGGACAGGAACCGGAGCAGCTCCAGTTATT 602572

Qy         87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArg 106
              |||   ||||||:::||||||||||||||||||||||||||||||||:::   ||||||:::
Db     602573 GCGCAAGTTGCAAAAGAGCTTGGTATTTTAACAGTTGGAGTTGTAACAAAGCCTTTTAAG 602632

Qy        107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyr 126
              |||||||||   :::   |||||||||      |||:::      |||   |||
Db     602633 TTTGAAGGTCCTAAGAAGTTGAGACTTGCTGAGCAGGGAATAAATAACTTAAGAAAGTCT 602692

Qy        127 ValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThr 146
              ||||||||||||||||:::::|||||||||  ||||||  :::  ||||||  |||
Db     602693 GTAGATACATTGATCATTATTCCAAATCAAAAGCTTTTAACTGTTGTTGACAAAAGAACC 602752

Qy        147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
              |||:::  :::||||||:::  ||||||||||||||      ||||||:::|||:::
Db     602753 ACCATTAAAGATGCTTTTAAGCGTGCAGATGATGTTCTTAGAATGGGCGTTCAAGGTATT 602812

Qy        167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::      |||   :::|||:::
Db     602813 GCAGGGCTTATTATTGAGCATGGAGAGGTTAATATT 602848

RESULT 9
US-09-830-228-1
; Sequence 1, Application US/09830228
;   GENERAL INFORMATION:
;        APPLICANT: Gil Choi et. al.
;        TITLE OF INVENTION: Borrelia burgdorferi Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 155
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/830,228
;             FILING DATE: 24-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: PCT/US98/12764
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brookes, A. Anders
;             REGISTRATION NUMBER: 36,373
;             REFERENCE/DOCKET NUMBER: PB370PCT
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 910715 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-830-228-1

Alignment Scores:
Pred. No.:              6.2e-40        Length:         910715
Score:                  506.00         Matches:        101
Percent Similarity:     74.42%         Conservative:   27
Best Local Similarity:  58.72%         Mismatches:     44
Query Match:            57.18%         Indels:         0
DB:                     32             Gaps:           0

US-09-770-509-2 (1-178) x US-09-830-228-1 (1-910715)

Qy          7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
              ||||||||||||||||||||||||||||  |||||||||      |||:::||||||
Db     602333 GTTGAATTTATTGTGGCTAATACCGATCTTCAGGCTCTCCAAACTTCTATTGCTCCCATA 602392

Qy         27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
              ||||||  ||||||      :::|||  ||||||||||||||  |||||||||:::
Db     602393 AAAATTGCCCTTGGAGCAAAAGTTACAGCAGGGCTTGGTGCTGGGGGAAAGCCTGAGATT 602452

Qy         47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
              |||:::  :::|||||:::            |||:::      |||      ::::::
Db     602453 GGACAAGCTGCAGCAGAGGAAGACATAGATGTTATACGAAATCATCTTTCTGGTGCCGAT 602512

Qy         67 MetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValVal 86
              |||:::|||||||||  ||||||||||||||||||  ||||||||||||||||||:::
Db     602513 ATGGTGTTTATTACTGCTGGTATGGGGGGCGGGACAGGAACCGGAGCAGCTCCAGTTATT 602572

Qy         87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArg 106
              |||   ||||||:::||||||||||||||||||||||||||||||||:::   ||||||:::
Db     602573 GCGCAAGTTGCAAAAGAGCTTGGTATTTTAACAGTTGGAGTTGTAACAAAGCCTTTTAAG 602632

Qy        107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyr 126
              |||||||||   :::   |||||||||      |||:::      |||   |||
Db     602633 TTTGAAGGTCCTAAGAAGTTGAGACTTGCTGAGCAGGGAATAAATAACTTAAGAAAGTCT 602692

Qy        127 ValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThr 146
              ||||||||||||||||:::::|||||||||  ||||||  :::  ||||||  |||
Db     602693 GTAGATACATTGATCATTATTCCAAATCAAAAGCTTTTAACTGTTGTTGACAAAAGAACC 602752

Qy        147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
              |||:::  :::||||||:::  ||||||||||||||      ||||||:::|||:::
Db     602753 ACCATTAAAGATGCTTTTAAGCGTGCAGATGATGTTCTTAGAATGGGCGTTCAAGGTATT 602812

Qy        167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::      |||   :::|||:::
Db     602813 GCAGGGCTTATTATTGAGCATGGAGAGGTTAATATT 602848

RESULT 10
```

```
US-09-663-779-1189/c
; Sequence 1189, Application US/09663779
; GENERAL INFORMATION:
;  APPLICANT:       Corbin, David R.
;  APPLICANT:  Malvar, Thomas M.
;  APPLICANT:  Shukla, Hridayabhiranjan
;  TITLE OF INVENTION:       BACILLUS THURINGIENSIS CHROMOSOMAL GENOME SEQUENCES AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE:       38-21(51376)B
;  CURRENT APPLICATION NUMBER: US/09/663,779
;  CURRENT FILING DATE:  2000-09-15
;  PRIOR APPLICATION NUMBER:       60/154,678
;  PRIOR FILING DATE:       1999-09-17
;  NUMBER OF SEQ ID NOS:       8283
; SEQ ID NO 1189
;   LENGTH: 7375
;   TYPE: DNA
;   ORGANISM: Bacillus thuringiensis
US-09-663-779-1189

Alignment Scores:
Pred. No.:              2.38e-42        Length:         7375
Score:                  505.00          Matches:        100
Percent Similarity:     74.29%          Conservative:   30
Best Local Similarity:  57.14%          Mismatches:     45
Query Match:            57.06%          Indels:         0
DB:                     26              Gaps:           0

US-09-770-509-2 (1-178) x US-09-663-779-1189 (1-7375)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              :::::|||||||:::||||||      |||||||||   |||||||||       |||
Db       1469 GTACAAGGTGTAGACTTTATCGCTGTGAACACTGATGCACAAGCATTAAATCTATCAAAA 1410

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||       |||:::   :::|||       :::|||:::||||||||||||||:::
Db       1409 GCTGAAACAAAAATGCAAATTGGTGGAAAATTAACGCGTGGACTTGGTGCAGGCGCAAAC 1350

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||::::::||||||:::   |||       ||||||      |||:::
Db       1349 CCTGAAGTAGGGAAAAAAGCTGCAGAAGAAAGTAAAGAACAGATCCAAGAAGCACTTCGT 1290

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                :::::|||:::|||:::|||   ||||||||||||||||||   ||||||||||||
Db       1289 GGTGCGGATATGGTCTTCGTAACTGCCGGTATGGGCGGTGGAACTGGAACTGGTGCAGCT 1230

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   ||||||:::||||||||   ||||||||||||||||||:::
Db       1229 CCAGTTGTTGCTCAAGTTGCAAAAGAATTAGGTGCTTTAACAGTTGGTGTTGTAACACGT 1170

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||       |||   :::|||:::
Db       1169 CCATTTACTTTTGAAGGACGTAAGCGTGCGACGCAAGCAGCATCTGGTATTGCAGCATTT 1110

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   |||||||||||||||||||:::|||||||      ||||||   :::   |||
Db       1109 AAAGAAAATGTAGATACTCTTATTGTAATTCCAAACGATCGCTTATTAGAGATTGTTGAT 1050

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db       1049 AAAAATACGCCAATGTTAGAAGCATTCCGTGAAGCTGATAACGTATTACGTCAAGGTGTT 990

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::|||||||||      ||||||||||||||||||
Db        989 CAAGGTATTTCGGATTTAATTGCAACGCCAGGTTTAATTAACTTA 945

RESULT 11
US-09-754-468-42
; Sequence 42, Application US/09754468
; GENERAL INFORMATION:
;  APPLICANT: Iversen, Patrick L.
;  TITLE OF INVENTION: Antisense Antibacterial Cell Division
;  TITLE OF INVENTION:  Composition and Method
;  FILE REFERENCE: 0450-0033.30
;  CURRENT APPLICATION NUMBER: US/09/754,468
;  CURRENT FILING DATE:  2001-01-04
;  PRIOR APPLICATION NUMBER: US 60/174,484
;  PRIOR FILING DATE: 2000-01-04
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
;   LENGTH: 1864
;   TYPE: DNA
;   ORGANISM: Bartonella henselae
US-09-754-468-42

Alignment Scores:
Pred. No.:              7.4e-43         Length:         1864
Score:                  503.00          Matches:        103
Percent Similarity:     73.71%          Conservative:   26
Best Local Similarity:  58.86%          Mismatches:     46
Query Match:            56.84%          Indels:         0
DB:                     29              Gaps:           0

US-09-770-509-2 (1-178) x US-09-754-468-42 (1-1864)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::||||||:::|||:::|||||||||||||||   |||||||||       |||
Db       223 CTTCAAGGAGTTGATTTTGTTGTTGCTAATACAGATGCGCAGGCTTTGGCTATGTCAAAG 282

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||          |||   ||||||      :::|||:::||||||||||||||:::
Db       283 GCTGAACGTGTTATCCAGCTTGGTGCAGCAGTTACAGAAGGTTTAGGTGCTGGTGCTTTA 342

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::   :::||||||:::          :::|||         |||
Db       343 CCGGAAGTTGGACAAGCGGCTGCAGAGGAATGTATTGATGAAATTATCGACCATCTCGCA 402

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ||||||:::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db       403 GATTCCCATATGATTTTCATTACTGCTGGTATGGGAGGAGGTACGGGAACAGGAGCAGCA 462

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||      |||||||||   ||||||||||||||||||||||||:::
Db       463 CCTGTTGTTGCACGTGCTGCTCGTGAAAAAGGTATTTTGACCGTTGGTGTTGTGACAAAG 522

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||:::   ||||||      |||   :::   |||   ||||||:::::||||||
Db       523 CCATTTCAGTTTGAAGGCGCTCGCCGTATGAAAACGGCAGAGGCTGGTATTGAAGAATTA 582

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 |||   ||||||||||||||||||:::||||||||||||||||      :::||||||
Db       583 CAAAAATCTGTTGATACATTGATTGTTATACCTAATCAGAATCTTTTCCGTATTGCAGAT 642

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              :::   ||||||      :::||||||      ||||||   ||||||      ||||||
Db       643 GAAAAAACAACCTTTGCTGATGCTTTTGCTATGGCTGACCAAGTGCTTTACTCTGGTGTT 702

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                    :::||||||||||:::::::::   ||||||||||||||
Db       703 GCTTCCATTACAGATCTCATGATTAAAGAAGGGCTCATTAACCTT 747

RESULT 12
US-09-974-300-180
; Sequence 180, Application US/09974300
; GENERAL INFORMATION:
;  APPLICANT: Berka, Randy M.
;  APPLICANT:  Clausen, Ib Groth
;  TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;  TITLE OF INVENTION:  Expression
;  FILE REFERENCE: 10085.500-US
;  CURRENT APPLICATION NUMBER: US/09/974,300
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/680,598
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: 60/279,526
```

```
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 8481
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 180
;   LENGTH: 1134
;   TYPE: DNA
;   ORGANISM: Bacillus licheniformis
US-09-974-300-180

Alignment Scores:
Pred. No.:                5.19e-43      Length:          1134
Score:                    502.00        Matches:         97
Percent Similarity:       74.01%        Conservative:    34
Best Local Similarity:    54.80%        Mismatches:      46
Query Match:              56.72%        Indels:          0
DB:                       36            Gaps:            0

US-09-770-509-2 (1-178) x US-09-974-300-180 (1-1134)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              :::   ::::::||||||||||||||||||      |||||||||   |||||||||
Db         97 AATGACGTTCAGGGAGTCGAGTTTATCGCAGTCAACACGGATGCTCAGGCTCTCAACCTG 156

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||   |||      |||:::   :::|||      :::|||:::||||||||||||||||
Db        157 TCAAAAGCGGAAACGAAAATGCAGATCGGTGCGAAGCTGACGCGCGGGCTCGGCGCCGGA 216

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   ||||||:::||||||:::::::||||||::   |||      |||:::
Db        217 GCCAATCCGGAAGTGGGCAAAAAAGCCGCAGAGGAAAGCAAAGAACAAATTGAAGAAGCG 276

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::   :::::::|||:::|||:::|||   ||||||||||||||||||   ||||||
Db        277 CTGAAAGGTGCAGATATGGTGTTCGTCACAGCCGGAATGGGCGGCGGAACAGGAACGGGC 336

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||:::|||   :::|||::::::|||||   ||||||||||||||||||
Db        337 GCGGCACCTGTCATCGCACAAATCGCAAAAGATCTGGGCGCATTGACTGTCGGCGTTGTC 396

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||      |||         |||      |||:::
Db        397 ACAAGGCCGTTTACCTTTGAAGGAAGAAAAAGACAGCTTCAGGCTGCAGGCGGTATTTCA 456

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              :::   :::  ||||||||||||||||||:::||||||      ||||||   :::
Db        457 GCAATGAAGGAAGCCGTCGACACCCTGATCGTCATTCCGAATGACCGTCTTCTTGAAATC 516

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              ||||||:::|||   ||||||||||||||||||   ||||||:::||||||   :::
Db        517 GTCGATAAAAACACACCGATGCTTGAAGCGTTCCGTGAAGCGGACAACGTTCTCCGCCAA 576

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||:::::::||||||||      ||||||||||||||||||
Db        577 GGTGTTCAGGGCATTTCAGACCTGATCGCAACGCCTGGACTGATCAACCTT 627

RESULT 13
US-60-045-649-854
; Sequence 854, Application US/60045649
;  GENERAL INFORMATION:
;    APPLICANT:  Lagace, Robert E.
;    APPLICANT:  Corely, Neil C.
;    APPLICANT:  Russo, Frank D.
;    APPLICANT:  Hann, Amy L.
;    TITLE OF INVENTION:  NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
;    TITLE OF INVENTION:  GENOME, FRAGMENTS THEREOF, AND USES THEREOF
;    NUMBER OF SEQUENCES:  1466
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/60/045,649
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PM-0003P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 845-4166
;  INFORMATION FOR SEQ ID NO:  854:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3993 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  EF1c0854
US-60-045-649-854

Alignment Scores:
Pred. No.:                1.03e-41      Length:          3993
Score:                    496.00        Matches:         100
Percent Similarity:       73.14%        Conservative:    28
Best Local Similarity:    57.14%        Mismatches:      47
Query Match:              56.05%        Indels:          0
DB:                       48            Gaps:            0

US-09-770-509-2 (1-178) x US-60-045-649-854 (1-3993)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||   ||||||||||||   |||||||||      |||
Db       2707 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 2766

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||         |||:::||||||||||||||||||:::
Db       2767 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 2826

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::||||||:::   :::   |||      ||||||
Db       2827 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 2886

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              :::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db       2887 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 2946

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db       2947 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 3006

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||      |||:::      |||
Db       3007 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCCGCTGAAGGAATTGCCTTATTA 3066

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::   ||||||||||||:::::::::   |||      ||||||   :::   |||
Db       3067 ANAGAAAACGTTGATACACTATTAATTATCTCAAACAACCGCTTATTAGAAGTCGTTGAC 3126

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::|||||
Db       3127 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 3186

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||   :::||||||
Db       3187 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTA 3231
```

```
RESULT 14
US-60-046-653-853
; Sequence 853, Application US/60046653
;  GENERAL INFORMATION:
;    APPLICANT:  Lagace, Robert E.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Russo, Frank D.
;    APPLICANT:  Hann, Amy L.
;    APPLICANT:  Heath, Joe D.
;    TITLE OF INVENTION:  NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
;    TITLE OF INVENTION:  GENOME, FRAGMENTS THEREOF, AND USES THEREOF
;    NUMBER OF SEQUENCES:  1449
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/60/046,653
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CERRONE, MICHAEL C.
;      REGISTRATION NUMBER:  39,132
;      REFERENCE/DOCKET NUMBER:  PM-0003-1 P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 855-0555
;      TELEFAX:  (415) 845-4166
;  INFORMATION FOR SEQ ID NO:  853:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3993 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  EFA1c853
US-60-046-653-853

Alignment Scores:
Pred. No.:              1.03e-41        Length:         3993
Score:                  496.00          Matches:        100
Percent Similarity:     73.14%          Conservative:   28
Best Local Similarity:  57.14%          Mismatches:     47
Query Match:            56.05%          Indels:         0
DB:                     48              Gaps:           0

US-09-770-509-2 (1-178) x US-60-046-653-853 (1-3993)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||   ||||||||||   |||||||||      |||
Db       2707 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 2766

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||          |||   ||||||         |||:::||||||||||||||||||:::
Db       2767 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 2826

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::||||||:::   :::      |||         ||||||
Db       2827 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 2886

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db       2887 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 2946

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              |||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db       2947 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 3006

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||      |||:::      |||
Db       3007 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCCGCTGAAGGAATTGCCTTATTA 3066

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::::::::   |||      ||||||   :::   |||
Db       3067 ANAGAAAACGTTGATACACTATTAATTATCTCAAACAACCGCTTATTAGAAGTCGTTGAC 3126

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db       3127 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 3186

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::::||||||||      ||||||   :::||||||
Db       3187 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTA 3231

RESULT 15
PCT-US02-03987-6542
; Sequence 6542, Application PC/TUS0203987
; GENERAL INFORMATION:
;  APPLICANT: Elitra Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
;  TITLE OF INVENTION:  Proliferation
;  FILE REFERENCE: ELITRA.028VPC
;  CURRENT APPLICATION NUMBER: PCT/US02/03987
;  CURRENT FILING DATE:  2002-02-02
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 15811
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6542
;   LENGTH: 1233
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(1233)
PCT-US02-03987-6542

Alignment Scores:
Pred. No.:              3.19e-42        Length:         1233
Score:                  495.00          Matches:        100
Percent Similarity:     73.14%          Conservative:   28
Best Local Similarity:  57.14%          Mismatches:     47
Query Match:            55.93%          Indels:         0
DB:                     1               Gaps:           0

US-09-770-509-2 (1-178) x PCT-US02-03987-6542 (1-1233)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||   ||||||||||   |||||||||      |||
Db        106 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 165

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||          |||   ||||||         |||:::||||||||||||||||||:::
Db        166 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 225

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::||||||:::   :::      |||         ||||||
Db        226 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 285

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        286 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 345

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              |||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db        346 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 405

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
```

```
                ||||||      |||||||||   |||   |||   |||      |||:::      |||
Db        406 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCCGCTGAAGGAATTGCCTTATTA 465

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                :::   ||||||||||||:::::::::   |||      ||||||   :::   |||
Db        466 AAAGAAAACGTTGATACACTATTAATTATCTCAAACAACCGCTTATTAGAAGTCGTTGAC 525

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
                |||   |||   ||||||||||||||||||   ||||||:::||||||   :::|||||||
Db        526 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 585

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                :::|||::::::|||||||||      ||||||   :::||||||
Db        586 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTA 630
```

Search completed: June 2, 2003, 11:06:15
Job time : 3441 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 09:41:39 ; Search time 740 Seconds
(without alignments)
1503.471 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7080287 seqs, 3125191874 residues

Total number of hits satisfying chosen parameters: 14160574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09770509/runat_27052003_083652_18961/app_query.fasta_1.327
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnpn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09770509_@CGN_1_1_671_@runat_27052003_083652_18961 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq2:*
7: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq3:*
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	510	57.6	1176	9	US-10-282-122A-24440	Sequence 24440, A
	2	510	57.6	213251	8	US-10-398-221-6	Sequence 6, Appli
c	3	510	57.6	3011208	8	US-10-398-221-2058	Sequence 2058, Ap
	4	507	57.3	1122	9	US-10-282-122A-15434	Sequence 15434, A
	5	506	57.2	1215	9	US-10-282-122A-10897	Sequence 10897, A
	6	505	57.1	1158	9	US-10-282-122A-10662	Sequence 10662, A
	7	505	57.1	1170	9	US-10-282-122A-9479	Sequence 9479, Ap
	8	505	57.1	1182	9	US-10-282-122A-17147	Sequence 17147, A
	9	495	55.9	1230	9	US-10-282-122A-6285	Sequence 6285, Ap
	10	495	55.9	1308	6	US-09-134-000C-1253	Sequence 1253, Ap
	11	495	55.9	1308	6	US-09-134-000C-1253	Sequence 1253, Ap
	12	494	55.8	1245	8	US-10-417-884-3641	Sequence 3641, Ap
	13	494	55.8	1245	9	US-10-282-122A-21408	Sequence 21408, A
	14	492	55.6	1161	9	US-10-282-122A-25784	Sequence 25784, A
	15	490	55.4	1107	9	US-10-282-122A-16538	Sequence 16538, A
	16	490	55.4	1260	1	PCT-US02-36122-83	Sequence 83, Appl
	17	490	55.4	1260	1	PCT-US02-36123-5201	Sequence 5201, Ap
c	18	490	55.4	1754382	1	PCT-US02-36123-6651	Sequence 6651, Ap
	19	486	54.9	1188	8	US-10-156-761-6101	Sequence 6101, Ap
	20	486	54.9	1233	9	US-10-282-122A-17831	Sequence 17831, A
	21	486	54.9	9025608	8	US-10-156-761-1	Sequence 1, Appli
	22	485	54.8	1137	9	US-10-282-122A-26150	Sequence 26150, A
	23	485	54.8	1140	9	US-10-282-122A-28464	Sequence 28464, A
	24	482	54.5	1179	9	US-10-282-122A-35322	Sequence 35322, A
	25	481	54.4	1140	9	US-10-282-122A-27607	Sequence 27607, A
	26	480	54.2	1173	6	US-09-950-084-1797	Sequence 1797, Ap
	27	480	54.2	1173	9	US-10-282-122A-7824	Sequence 7824, Ap
	28	480	54.2	1182	9	US-10-282-122A-34548	Sequence 34548, A
	29	480	54.2	1185	9	US-10-092-411A-1997	Sequence 1997, Ap
	30	480	54.2	2893	10	US-60-009-861-190	Sequence 190, App
	31	480	54.2	3931	9	US-10-329-624-342	Sequence 342, App
	32	480	54.2	98958	6	US-09-950-084-7442	Sequence 7442, Ap
	33	472	53.3	1260	9	US-10-282-122A-37929	Sequence 37929, A
	34	471	53.2	1194	9	US-10-282-122A-31723	Sequence 31723, A
	35	468	52.9	1185	9	US-10-282-122A-33518	Sequence 33518, A
c	36	466	52.7	29112	1	PCT-US02-32727-104	Sequence 104, App
c	37	466	52.7	29112	6	US-09-978-825-104	Sequence 104, App
c	38	466	52.7	29112	9	US-10-057-498-104	Sequence 104, App
	39	465	52.5	1257	9	US-10-282-122A-40313	Sequence 40313, A
	40	464	52.4	1185	9	US-10-282-122A-30463	Sequence 30463, A
	41	464	52.4	1203	8	US-10-419-128-7557	Sequence 7557, Ap
	42	464	52.4	1203	9	US-10-366-683-7557	Sequence 7557, Ap
c	43	464	52.4	1314	8	US-10-419-128-7784	Sequence 7784, Ap
c	44	464	52.4	1314	9	US-10-366-683-7784	Sequence 7784, Ap
	45	464	52.4	1320	9	US-10-282-122A-38466	Sequence 38466, A

ALIGNMENTS

RESULT 1
US-10-282-122A-24440
; Sequence 24440, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24440
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Listeria monocytogenes
US-10-282-122A-24440

Alignment Scores:

Pred. No.:	2.3e-48	Length:	1176
Score:	510.00	Matches:	98
Percent Similarity:	76.00%	Conservative:	35
Best Local Similarity:	56.00%	Mismatches:	42
Query Match:	57.63%	Indels:	0
DB:	9	Gaps:	0

US-09-770-509-2 (1-178) x US-10-282-122A-24440 (1-1176)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||      |||||||||   |||||||||      :::
Db        103 GTTCAAGGAGTAGAATTTATCTCCGTTAATACAGACGCTCAAGCACTTAACTTAGCAAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||       |||:::   :::|||      :::|||:::|||||||||||||||||:::
Db        163 GCAGAAACAAAATTACAAATCGGTACAAAATTAACGCGTGGTTTAGGTGCGGGAGCTGTA 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||||||:::::|||||||:::   :::      |||:::      |||:::
Db        223 CCTGAAATTGGTAAAAAAGCTGCAGAAGAAAGCCGCGAACAAATTGAAGAAGCTTTAAAA 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||:::|||:::|||:::|||   |||||||||||||||||   ||||||||||||
Db        283 GGCTCTGATATGGTATTCGTAACTGCTGGAATGGGCGGCGGAACTGGAACTGGGGCTGCA 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   :::|||:::|||:::|||   |||||||||||||||||:::
Db        343 CCTGTTATCGCTCAAATCGCAAAAGAAATGGGCGCTTTAACAGTAGGTGTTGTTACTCGA 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   ||||||:::   |||      |||   :::   :::
Db        403 CCATTTGGTTTTGAAGGACCAAAACGTACGAAACAAGCCTTAACTGGAACAGAAGCAATG 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   |||||||||||||||||:::||||||      ||||||   :::   |||
Db        463 AAAGAAGCGGTGGATACGTTAATTGTTATTCCTAATGACCGTTTACTTCAAATTGTGGAT 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   |||||||||||||||||   ||||||:::|||||||   :::|||||||
Db        523 AAAAATACACCGATGCTTGAAGCTTTCCGTGAAGCAGATAATGTTTTACGTCAAGGTGTA 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||||||||||||||
Db        583 CAAGGGATTTCTGATTTGATTGCCGTTCCTGGTTTAATTAACTTA 627
```

RESULT 2
US-10-398-221-6
; Sequence 6, Application US/10398221
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 213251
; TYPE: DNA
; ORGANISM: Listeria innocua
US-10-398-221-6

Alignment Scores:

Pred. No.:	2.78e-45	Length:	213251
Score:	510.00	Matches:	98
Percent Similarity:	76.00%	Conservative:	35
Best Local Similarity:	56.00%	Mismatches:	42
Query Match:	57.63%	Indels:	0
DB:	8	Gaps:	0

US-09-770-509-2 (1-178) x US-10-398-221-6 (1-213251)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||      |||||||||   |||||||||      :::
Db     170157 GTTCAAGGAGTAGAATTTATCTCCGTTAATACAGACGCTCAAGCACTTAATTTAGCAAAA 170216

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||       |||:::   :::|||      :::|||:::|||||||||||||||||:::
Db     170217 GCAGAAACAAAATTACAAATCGGTACAAAATTAACGCGTGGTTTAGGCGCGGGTGCTGTA 170276

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||||||:::::|||||||:::   :::      |||:::      |||:::
Db     170277 CCTGAAATTGGTAAAAAAGCAGCAGAAGAAAGTCGCGAACAAATTGAAGAAGCTTTAAAA 170336

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||:::|||:::|||:::|||   |||||||||||||||||   ||||||||||||
Db     170337 GGCTCTGATATGGTATTCGTAACTGCTGGGATGGGCGGCGGAACTGGAACTGGTGCTGCA 170396

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   :::|||:::|||:::|||   |||||||||||||||||:::
Db     170397 CCTGTTATCGCTCAAATCGCAAAAGAAATGGGCGCTCTAACTGTCGGTGTTGTTACACGA 170456

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   ||||||:::   |||      |||   :::   :::
Db     170457 CCATTTGGTTTTGAAGGACCAAAACGTACGAAACAAGCCCTAACTGGAACAGAAGCAATG 170516

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   |||||||||||||||||:::||||||      ||||||   :::   |||
Db     170517 AAAGAAGCGGTGGATACTTTAATTGTTATCCCTAACGACCGTTTACTTCAAATTGTTGAT 170576

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   |||||||||||||||||   ||||||:::|||||||   :::|||||||
Db     170577 AAAAATACGCCAATGCTTGAAGCTTTCCGTGAAGCAGATAATGTTTTACGTCAAGGGGTA 170636

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||||||||||||||
Db     170637 CAAGGGATTTCTGATTTGATTGCCGTTCCTGGTTTAATTAACTTA 170681
```

RESULT 3
US-10-398-221-2058/c
; Sequence 2058, Application US/10398221
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04

```
;  NUMBER OF SEQ ID NOS: 4025
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2058
;   LENGTH: 3011208
;   TYPE: DNA
;   ORGANISM: Listeria innocua
US-10-398-221-2058

Alignment Scores:
Pred. No.:                1.03e-43      Length:          3011208
Score:                    510.00        Matches:         98
Percent Similarity:       76.00%        Conservative:    35
Best Local Similarity:    56.00%        Mismatches:      42
Query Match:              57.63%        Indels:          0
DB:                       8             Gaps:            0

US-09-770-509-2 (1-178) x US-10-398-221-2058 (1-3011208)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              :::::|||||||||||||||||      |||||||||   |||||||||      :::
Db    2161845 GTTCAAGGAGTAGAATTTATCTCCGTTAATACAGACGCTCAAGCACTTAATTTAGCAAAA 2161786

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||      :::|||:::||||||||||||||||:::
Db    2161785 GCAGAAACAAAATTACAAATCGGTACAAAATTAACGCGTGGTTTAGGCGCGGGTGCTGTA 2161726

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||||||:::::|||||:::   :::      |||:::      |||:::
Db    2161725 CCTGAAATTGGTAAAAAAGCAGCAGAAGAAAGTCGCGAACAAATTGAAGAAGCTTTAAAA 2161666

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||:::|||:::|||:::|||   ||||||||||||||||||   ||||||||||||
Db    2161665 GGCTCTGATATGGTATTCGTAACTGCTGGGATGGGCGGCGGAACTGGAACTGGTGCTGCA 2161606

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   :::|||:::|||:::|||   ||||||||||||||||||:::
Db    2161605 CCTGTTATCGCTCAAATCGCAAAAGAAATGGGCGCTCTAACTGTCGGTGTTGTTACACGA 2161546

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   ||||||:::   |||      |||   :::   :::
Db    2161545 CCATTTGGTTTTGAAGGACCAAAACGTACGAAACAAGCCCTAACTGGAACAGAAGCAATG 2161486

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||||||||:::|||||||      ||||||   :::   |||
Db    2161485 AAAGAAGCGGTGGATACTTTAATTGTTATCCCTAACGACCGTTTACTTCAAATTGTTGAT 2161426

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db    2161425 AAAAATACGCCAATGCTTGAAGCTTTCCGTGAAGCAGATAATGTTTTACGTCAAGGGGTA 2161366

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::::|||||||||      ||||||||||||||||||
Db    2161365 CAAGGGATTTCTGATTTGATTGCCGTTCCTGGTTTAATTAACTTA 2161321

RESULT 4
US-10-282-122A-15434
; Sequence 15434, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15434
;   LENGTH: 1122
;   TYPE: DNA
;   ORGANISM: Clostridium acetobutylicum
US-10-282-122A-15434

Alignment Scores:
Pred. No.:                4.77e-48      Length:          1122
Score:                    507.00        Matches:         98
Percent Similarity:       76.00%        Conservative:    35
Best Local Similarity:    56.00%        Mismatches:      42
Query Match:              57.29%        Indels:          0
DB:                       9             Gaps:            0

US-09-770-509-2 (1-178) x US-10-282-122A-15434 (1-1122)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::   ||||||||||||      |||||||||   |||||||||      |||
Db        103 CTTAAGAATGTGGAATTTATTGCCATAAACACTGATAAGCAGGCATTGGCTCTTTCGCAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      ||||||   :::|||      :::|||||||||||||||||||:::
Db        163 GCATCACAGAAGATACAAATAGGTGATAAGCTTACTAAAGGGTTAGGAGCAGGAGCTAAT 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::::::||||||:::   |||   :::|||   :::   ::::::
Db        223 CCTGAAATAGGTCAAAAGGCTGCTGAAGAAAGTAAGGATGAAATATCTCAGGCTATAAAA 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        283 GGTGCTGATATGGTATTTATAACTGCTGGTATGGGTGGTGGAACTGGAACTGGTGCAGCA 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::   :::|||||||||||||||||||||||:::
Db        343 CCTGTCGTTGCAGAAATAGCTAAATCCATGGGAATATTAACAGTTGGAGTTGTTACAAAG 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||   :::|||:::|||   |||
Db        403 CCTTTCCCTTTTGAAGGAAGAAAGAGAATGCTCCATGCTGAGTCAGGAATTAAAACACTA 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::   :::|||||||:::   ||||||||||:::   |||
Db        463 AAAGAAAGAGTTGATACTTTAGTTACAATTCCAAATGAAAGATTGCTTGCTATTGTAGAT 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   ||||||:::::::||||||||:::   ||||||||||||||   :::||||||
Db        523 AAAAAGACAACTTTAGTAGAAGCCTTTAAATCAGCAGATGATGTTTTAAGACAAGGTGTT 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
```

```
                ::::|||::::::|||||||||      |||||||||:::|||||||
Db        583 CAAGGTATATCAGATTTGATAACTATTCCTGGTCTTGTAAACCTT 627

RESULT 5
US-10-282-122A-10897
; Sequence 10897, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10897
;   LENGTH: 1215
;   TYPE: DNA
;   ORGANISM: Borrelia burgdorferi
US-10-282-122A-10897

Alignment Scores:
Pred. No.:              6.92e-48        Length:         1215
Score:                  506.00          Matches:        101
Percent Similarity:     74.42%          Conservative:   27
Best Local Similarity:  58.72%          Mismatches:     44
Query Match:            57.18%          Indels:         0
DB:                     9               Gaps:           0

US-09-770-509-2 (1-178) x US-10-282-122A-10897 (1-1215)

Qy          7 ValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHis 26
              ||||||||||||||||||||||||||  |||||||||      |||:::||||||
Db        154 GTTGAATTTATTGTGGCTAATACCGATCTTCAGGCTCTCCAAACTTCTATTGCTCCCATA 213

Qy         27 LysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeu 46
              ||||||   ||||||      :::|||  ||||||||||||||   |||||||||:::
Db        214 AAAATTGCCCTTGGAGCAAAAGTTACAGCAGGGCTTGGTGCTGGGGGAAAGCCTGAGATT 273

Qy         47 GlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsn 66
              |||:::   :::|||||||:::            |||:::      |||      ::::::
Db        274 GGACAAGCTGCAGCAGAGGAAGACATAGATGTTATACGAAATCATCTTTCTGGTGCCGAT 333

Qy         67 MetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValVal 86
              |||:::|||||||||   ||||||||||||||||||   ||||||||||||||||||:::
Db        334 ATGGTGTTTATTACTGCTGGTATGGGGGGCGGGACAGGAACCGGAGCAGCTCCAGTTATT 393

Qy         87 AlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPheArg 106
              |||   ||||||:::||||||||||||||||||||||||||||||:::   ||||||:::
Db        394 GCGCAAGTTGCAAAAGAGCTTGGTATTTTAACAGTTGGAGTTGTAACAAAGCCTTTTAAG 453

Qy        107 SerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyr 126
                 |||||||||   :::   |||||||||      |||:::      |||   |||
Db        454 TTTGAAGGTCCTAAGAAGTTGAGACTTGCTGAGCAGGGAATAAATAACTTAAGAAAGTCT 513

Qy        127 ValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThr 146
              ||||||||||||||||:::::::|||||||||   ||||||   :::   ||||||   |||
Db        514 GTAGATACATTGATCATTATTCCAAATCAAAAGCTTTTAACTGTTGTTGACAAAAGAACC 573

Qy        147 ThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyVal 166
              |||:::   :::|||||||:::   ||||||||||||||||      ||||||:::|||:::
Db        574 ACCATTAAAGATGCTTTTAAGCGTGCAGATGATGTTCTTAGAATGGGCGTTCAAGGTATT 633

Qy        167 ThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
                 ||||||:::      |||   :::|||:::
Db        634 GCAGGGCTTATTATTGAGCATGGAGAGGTTAATATT 669

RESULT 6
US-10-282-122A-10662
; Sequence 10662, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10662
;   LENGTH: 1158
;   TYPE: DNA
;   ORGANISM: Bacillus anthracis
US-10-282-122A-10662
```

Alignment Scores:

Pred. No.:	8.45e-48	Length:	1158
Score:	505.00	Matches:	100
Percent Similarity:	74.29%	Conservative:	30
Best Local Similarity:	57.14%	Mismatches:	45
Query Match:	57.06%	Indels:	0
DB:	9	Gaps:	0

US-09-770-509-2 (1-178) x US-10-282-122A-10662 (1-1158)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::||||||:::|||||      |||||||||  |||||||||     |||
Db        103 GTACAAGGTGTAGACTTTATCGCTGTGAATACTGATGCACAAGCATTAAATCTATCAAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||      :::|||:::||||||||||||||:::
Db        163 GCTGAAACGAAAATGCAAATTGGTGGAAAATTAACACGTGGACTTGGTGCAGGTGCAAAC 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||::::::||||||:::   |||      ||||||      |||:::
Db        223 CCTGAAGTAGGGAAAAAAGCTGCGGAAGAAAGTAAAGAACAGATCCAAGAAGCACTTCGT 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||:::|||   ||||||||||||||||   ||||||||||||
Db        283 GGTGCAGATATGGTCTTCGTAACTGCTGGTATGGGCGGCGGAACTGGAACTGGTGCAGCT 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   ||||||:::||||||||||   ||||||||||||||||:::
Db        343 CCAGTTGTTGCTCAAGTGGCAAAAGAACTAGGTGCATTAACAGTTGGTGTTGTAACACGT 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||   :::|||:::
Db        403 CCATTTACATTTGAAGGACGTAAGCGTGCAACACAAGCGGCATCTGGTATTGCGGCATTT 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::   ||||||||||||||||:::||||||      ||||||   :::   |||
Db        463 AAAGAAAATGTAGATACACTTATTGTAATTCCAAACGATCGCTTATTAGAGATTGTTGAT 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   ||||||||||||||||   ||||||:::||||||   :::||||||
Db        523 AAAAACACGCCAATGTTAGAGGCATTCCGTGAAGCGGATAACGTATTACGTCAAGGTGTT 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::||||||||      ||||||||||||||||||
Db        583 CAAGGTATTTCTGATTTAATTGCAACACCAGGTTTAATTAACTTA 627
```

RESULT 7
US-10-282-122A-9479
; Sequence 9479, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9479
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Bacillus anthracis
US-10-282-122A-9479

Alignment Scores:

Pred. No.:	8.57e-48	Length:	1170
Score:	505.00	Matches:	100
Percent Similarity:	74.29%	Conservative:	30
Best Local Similarity:	57.14%	Mismatches:	45
Query Match:	57.06%	Indels:	0
DB:	9	Gaps:	0

US-09-770-509-2 (1-178) x US-10-282-122A-9479 (1-1170)

```
Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::||||||:::|||||      |||||||||  |||||||||     |||
Db        103 GTACAAGGTGTAGACTTTATCGCTGTGAATACTGATGCACAAGCATTAAATCTATCAAAA 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      |||:::   :::|||      :::|||:::||||||||||||||:::
Db        163 GCTGAAACGAAAATGCAAATTGGTGGAAAATTAACACGTGGACTTGGTGCAGGTGCAAAC 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::||||||::::::||||||:::   |||      ||||||      |||:::
Db        223 CCTGAAGTAGGGAAAAAAGCTGCGGAAGAAAGTAAAGAACAGATCCAAGAAGCACTTCGT 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||:::|||   ||||||||||||||||   ||||||||||||
Db        283 GGTGCAGATATGGTCTTCGTAACTGCTGGTATGGGCGGCGGAACTGGAACTGGTGCAGCT 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   ||||||:::||||||||||   ||||||||||||||||:::
Db        343 CCAGTTGTTGCTCAAGTGGCAAAAGAACTAGGTGCATTAACAGTTGGTGTTGTAACACGT 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||   :::|||:::
Db        403 CCATTTACATTTGAAGGACGTAAGCGTGCAACACAAGCGGCATCTGGTATTGCGGCATTT 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::   ||||||||||||||||:::||||||      ||||||   :::   |||
Db        463 AAAGAAAATGTAGATACACTTATTGTAATTCCAAACGATCGCTTATTAGAGATTGTTGAT 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||:::|||   ||||||||||||||||   ||||||:::||||||   :::||||||
Db        523 AAAAACACGCCAATGTTAGAGGCATTCCGTGAAGCGGATAACGTATTACGTCAAGGTGTT 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::||||||||      ||||||||||||||||||
Db        583 CAAGGTATTTCTGATTTAATTGCAACACCAGGTTTAATTAACTTA 627
```

RESULT 8
US-10-282-122A-17147
; Sequence 17147, Application US/10282122A
; GENERAL INFORMATION:

```
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17147
;   LENGTH: 1182
;   TYPE: DNA
;   ORGANISM: Clostridium difficile
US-10-282-122A-17147

Alignment Scores:
Pred. No.:              8.69e-48        Length:         1182
Score:                  505.00          Matches:        99
Percent Similarity:     75.71%          Conservative:   35
Best Local Similarity:  55.93%          Mismatches:     43
Query Match:            57.06%          Indels:         0
DB:                     9               Gaps:           0

US-09-770-509-2 (1-178) x US-10-282-122A-17147 (1-1182)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              :::|||||:::|||||||||||||||||      |||||||||   |||||||||
Db        124 GCTCAACTAAAAGGCGTTGAATTTATATCAGTAAATACAGATAAACAAGCTTTATATACA 183

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||   |||   :::|||:::   :::|||:::   :::|||:::|||||||||||||||
Db        184 TCAAAAGCTGAATATAAAGTTCAAATAGGAGAAAAATTAACTAGAGGACTTGGAGCAGGA 243

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
              :::   ||||||:::||||||||||:::||||||:::   |||   :::|||   ::::::
Db        244 GCAAATCCAGAAGTAGGGAAAAGAGCTGCTGAGGAGAGTAAAGATGAAATAGTAAAGTTA 303

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              ||||||   ::::::|||:::|||:::|||   ||||||||||||||||||   ||||||
Db        304 CTTCAAGGAGCTGATATGGTCTTTGTAACAGCCGGAATGGGCGGTGGTACTGGTACAGGA 363

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||||||||   :::|||:::|||:::|||||||||||||||||||||||||

Db        364 GCTGCTCCAGTGGTTGCAGGCTTGGCAAAAGAAATGGGGATACTTACAGTAGGTGTAGTT 423

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||      |||   :::   |||      |||:::
Db        424 ACCAAACCTTTTGCATTTGAAGGTAAGATAAGAATGAAGAACGCTGAAGGTGGAATAGCA 483

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              ||||||          ||||||||||||||||   :::||||||      ||||||   :::
Db        484 GAACTTAAATCAAAAGTAGATACTCTTATAACGATACCAAATGATAGACTTTTACAGATA 543

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
                    |||:::|||:::||||||:::||||||      ||||||||||||||   :::
Db        544 GTACAAAAAAATACATCAATGTTAGATGCTTTTGCAGTTGCAGATGATGTATTAAAACAA 603

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              |||::::::   ::::::|||||||||         ||||||||||||||
Db        604 GGTATACAATCAATTTCAGACCTTATAGCAGTTGAAGGGTTAATAAACTTG 654

RESULT 9
US-10-282-122A-6285
; Sequence 6285, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6285
;   LENGTH: 1230
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
US-10-282-122A-6285

Alignment Scores:
Pred. No.:              1.29e-46        Length:         1230
Score:                  495.00          Matches:        100
Percent Similarity:     73.14%          Conservative:   28
Best Local Similarity:  57.14%          Mismatches:     47
Query Match:            55.93%          Indels:         0
```

```
DB:                 9           Gaps:          0

US-09-770-509-2 (1-178) x US-10-282-122A-6285 (1-1230)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||   ||||||||||||   ||||||||      |||
Db        106 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 165

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||         |||:::|||||||||||||||||||:::
Db        166 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 225

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||:::   :::      |||         ||||||
Db        226 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 285

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        286 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 345

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db        346 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 405

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||       |||:::      |||
Db        406 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCTGCTGAAGGAATTGCCTTATTA 465

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::::::::   |||       ||||||   :::   |||
Db        466 AAAGAAAACGTTGATACACTATTAATTATCTCAAATAACCGCTTATTAGAAGTCGTTGAC 525

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db        526 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 585

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||   :::|||||||
Db        586 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTG 630

RESULT 10
US-09-134-000C-1253
; Sequence 1253, Application US/09134000C
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1253
;   LENGTH: 1308
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-1253

Alignment Scores:
Pred. No.:              1.4e-46       Length:          1308
Score:                  495.00        Matches:         100
Percent Similarity:     73.14%        Conservative:    28
Best Local Similarity:  57.14%        Mismatches:      47
Query Match:            55.93%        Indels:          0
DB:                     6             Gaps:            0

US-09-770-509-2 (1-178) x US-09-134-000C-1253 (1-1308)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||   ||||||||||||   ||||||||      |||
Db        181 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 240

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||         |||:::|||||||||||||||||||:::
Db        241 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 300

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||:::   :::      |||         ||||||
Db        301 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 360

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 ::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        361 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 420

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db        421 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 480

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||       |||:::      |||
Db        481 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCTGCTGAAGGAATTGCCTTATTA 540

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::::::::   |||       ||||||   :::   |||
Db        541 AAAGAAAACGTTGATACACTATTAATTATCTCAAATAACCGCTTATTAGAAGTCGTTGAC 600

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db        601 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 660

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||   :::|||||||
Db        661 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTG 705

RESULT 11
US-09-134-000C-1253
; Sequence 1253, Application US/09134000C
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1253
;   LENGTH: 1308
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-1253

Alignment Scores:
Pred. No.:              1.4e-46       Length:          1308
Score:                  495.00        Matches:         100
Percent Similarity:     73.14%        Conservative:    28
Best Local Similarity:  57.14%        Mismatches:      47
Query Match:            55.93%        Indels:          0
DB:                     6             Gaps:            0

US-09-770-509-2 (1-178) x US-09-134-000C-1253 (1-1308)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||   ||||||||||||   ||||||||      |||
Db        181 GTTAAAGGCGTGGAATTTATCACAGCCAATACAGACGTTCAAGCATTAAAACATTCAAAA 240

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||         |||   ||||||         |||:::|||||||||||||||||||:::
Db        241 GCAGAAACAGTGATTCAATTAGGCCCTAAATACACTCGTGGTTTAGGTGCCGGTTCACAA 300

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::|||||:::   :::      |||         ||||||
```

```
Db        301 CCTGAAGTTGGCCAAAAAGCTGCAGAAGAAAGTGAACAAGTGATTTCAGAATCATTACAA 360

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              :::::::|||:::|||||||||   |||||||||||||||||   ||||||||||||
Db        361 GGCGCGGATATGATTTTCATTACTGCTGGTATGGGTGGCGGAACTGGTACAGGTGCTGCG 420

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||||||||   ||||||||||||||||||:::
Db        421 CCAGTAGTTGCAAAAATCGCTAAAGAATTAGGCGCTTTAACAGTTGGTGTAGTAACTCGT 480

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||       |||:::       |||
Db        481 CCATTTAGTTTTGAAGGTCCAAAACGTGGTCGTTTTGCTGCTGAAGGAATTGCCTTATTA 540

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::::::::   |||       ||||||   :::   |||
Db        541 AAAGAAAACGTTGATACACTATTAATTATCTCAAATAACCGCTTATTAGAAGTCGTTGAC 600

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||   |||||||:::||||||   :::||||||
Db        601 AAGAAAACGCCAATGCTTGAAGCATTTAGAGAAGCTGATAATGTATTACGTCAAGGTGTT 660

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||   :::||||||
Db        661 CAAGGGATTTCAGATTTAATCACTGCACCAGGTTACGTAAACTTG 705

RESULT 12
US-10-417-884-3641
; Sequence 3641, Application US/10417884
;   GENERAL INFORMATION:
;        APPLICANT: Lynn A Doucette-Stamm  and David Bush
;        TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                            ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;        NUMBER OF SEQUENCES: 7310
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;             STREET: 100 Beaver Street
;             CITY: Waltham
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02354
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: CD/ROM ISO9660
;             COMPUTER: PC
;             OPERATING SYSTEM: <Unknown>
;             SOFTWARE: ASCII
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/417,884
;             FILING DATE: 17-Apr-2003
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/107,532A
;             FILING DATE: 30-Jun-1998
;             APPLICATION NUMBER: 60/085,598
;             FILING DATE: 14 May 1998
;             APPLICATION NUMBER: 60/051571
;             FILING DATE: July 2, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Ariniello, Pamela Deneke
;             REGISTRATION NUMBER: 40,489
;             REFERENCE/DOCKET NUMBER: GTC-012
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (781)893-5007
;             TELEFAX: (781)893-8277
;   INFORMATION FOR SEQ ID NO: 3641:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1245 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: circular
;        MOLECULE TYPE: DNA (genomic)
;        HYPOTHETICAL: NO
;        ANTI-SENSE: NO
;        ORIGINAL SOURCE:
;             ORGANISM: Enterococcus faecium
;        FEATURE:
;             NAME/KEY:  misc_feature
;             LOCATION: (B) LOCATION 1...1245
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3641:
US-10-417-884-3641

Alignment Scores:
Pred. No.:                1.71e-46      Length:          1245
Score:                    494.00        Matches:         97
Percent Similarity:       73.14%        Conservative:    31
Best Local Similarity:    55.43%        Mismatches:      47
Query Match:              55.82%        Indels:          0
DB:                       8             Gaps:            0

US-09-770-509-2 (1-178) x US-10-417-884-3641 (1-1245)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||   ||||||||||||   |||||||||      |||
Db        109 GTTAAAGGTGTTGAATTCATTACAGCCAACACAGACGTGCAAGCATTAAAAAATTCAAAA 168

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||          |||   ||||||         |||:::|||||||||||||||||||:::
Db        169 GCTGAAACAGTTATCCAATTAGGCCCCAAATATACTCGTGGTTTAGGTGCTGGTTCTCAG 228

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::||||||:::   :::      ::::::      |||
Db        229 CCAGAAGTTGGACAAAAAGCGGCAGAAGAAAGTGAACAATCCTTACGCGAAGCTTTAGAC 288

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||||||||   |||||||||||||||||   ||||||||||||
Db        289 GGTGCAGATATGATTTTCATCACTGCAGGAATGGGTGGAGGAACCGGTACAGGTGCTGCA 348

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              |||:::|||||   :::|||:::|||||||||   |||||||||||||||||||:::
Db        349 CCAATCGTTGCAGGTATCGCGAAAGAATTGGGCGCATTGACTGTTGGTGTCGTGACACGT 408

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||       |||:::       |||
Db        409 CCATTCACTTTTGAAGGACCGAAACGTGGCCGTTTTGCTGCAGAAGGTATCGCTCGTTTG 468

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::::::::   |||       ||||||   :::   |||
Db        469 AAAGAAAATGTGGATACATTATTAATTATCTCAAATAACCGTTTATTAGAAGTTGTAGAT 528

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||   |||||||:::|||||   :::|||||||
Db        529 AAAAAGACACCGATGCTTGAAGCTTTCCGTGAAGCAGATAATGTTTTACGACAAGGTGTA 588

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||||||||      ||||||   :::||||||
Db        589 CAAGGTATCTCTGATTTGATCACAGCGCCAGGTTACGTTAACCTT 633

RESULT 13
US-10-282-122A-21408
; Sequence 21408, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
```

```
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21408
;   LENGTH: 1245
;   TYPE: DNA
;   ORGANISM: Enterococcus faecium
US-10-282-122A-21408

Alignment Scores:
Pred. No.:              1.71e-46        Length:         1245
Score:                  494.00          Matches:        97
Percent Similarity:     73.14%          Conservative:   31
Best Local Similarity:  55.43%          Mismatches:     47
Query Match:            55.82%          Indels:         0
DB:                     9               Gaps:           0

US-09-770-509-2 (1-178) x US-10-282-122A-21408 (1-1245)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              ::::::|||||||||||||||||   ||||||||||||   |||||||||      |||
Db        106 GTTAAAGGTGTTGAATTCATTACAGCCAACACAGACGTGCAAGCATTAAAAAATTCAAAA 165

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||          |||   ||||||         |||:::|||||||||||||||||:::
Db        166 GCTGAAACAGTTATCCAATTAGGCCCCAAATATACTCGTGGTTTAGGTGCTGGTTCTCAG 225

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              |||||||:::|||:::::::::||||||:::   :::     ::::::      |||
Db        226 CCAGAAGTTGGACAAAAAGCGGCAGAAGAAAGTGAACAATCCTTACGCGAAGCTTTAGAC 285

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        286 GGTGCAGATATGATTTTCATCACTGCAGGAATGGGTGGAGGAACCGGTACAGGTGCTGCA 345

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              |||:::||||||   :::|||:::||||||||||   |||||||||||||||||||:::
Db        346 CCAATCGTTGCAGGTATCGCGAAAGAATTGGGCGCATTGACTGTTGGTGTCGTGACACGT 405

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      |||||||||   |||   |||   |||       |||:::      |||
Db        406 CCATTCACTTTTGAAGGACCGAAACGTGGCCGTTTTGCTGCAGAAGGTATCGCTCGTTTG 465

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||:::::::::   |||       ||||||   :::   |||
Db        466 AAAGAAAATGTGGATACATTATTAATTATCTCAAATAACCGTTTATTAGAAGTTGTAGAT 525

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||   ||||||||||||||||||   ||||||:::||||||   :::||||||
Db        526 AAAAAGACACCGATGCTTGAAGCTTTCCGTGAAGCAGATAATGTTTTACGACAAGGTGTA 585

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::::|||||||||      ||||||   :::||||||
Db        586 CAAGGTATCTCTGATTTGATCACAGCGCCAGGTTACGTTAACCTT 630

RESULT 14
US-10-282-122A-25784
; Sequence 25784, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25784
;   LENGTH: 1161
;   TYPE: DNA
;   ORGANISM: Mycobacterium avium
US-10-282-122A-25784

Alignment Scores:
Pred. No.:              2.64e-46        Length:         1161
Score:                  492.00          Matches:        96
Percent Similarity:     74.86%          Conservative:   35
Best Local Similarity:  54.86%          Mismatches:     44
Query Match:            55.59%          Indels:         0
DB:                     9               Gaps:           0

US-09-770-509-2 (1-178) x US-10-282-122A-25784 (1-1161)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::|||||||||||||||||       |||||||||   |||||||||       |||
Db         97 CTCAAGGGCGTGGAGTTCATCGCGATCAACACCGACGCGCAAGCGCTGTTGATGAGCGAT 156

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||       |||:::   :::|||:::|||   |||:::||||||||||||||:::
Db        157 GCCGACGTCAAACTCGACGTCGGGCGCGACTCCACCCGTGGACTGGGCGCCGGCGCGGAC 216

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              ||||||:::|||:::::::::||||||      |||   :::|||:::   :::|||:::
Db        217 CCCGAGGTCGGCCGCAAGGCCGCCGAGGACGCCAAGGACGAGATCGAGGAGCTGCTGCGC 276
```

```
Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||:::|||   |||   ||||||||||||   ||||||   |||
Db        277 GGCGCCGACATGGTGTTCGTCACCGCCGGCGAGGGCGGCGGCACCGGGACCGGCGGGGCG 336

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              |||||||||||||||:::||||||:::||||||   ||||||||||||||||||:::
Db        337 CCCGTCGTCGCCAGCATCGCGCGCAAGCTCGGCGCTTTGACCGTGGGCGTGGTCACCCGG 396

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||:::      |||   ||||||:::      |||
Db        397 CCGTTCTCGTTCGAGGGCAAGCGGCGCAGCAACCAGGCCGAGGCGGGCATCAACGCGCTG 456

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::      ||||||||||||||||:::||||||      ||||||   :::   |||
Db        457 CGGGAGAGCTGCGACACCCTGATCGTCATCCCCAACGACCGGCTGCTGCAGATGGGCGAC 516

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              :::   ::::::::::::|||||||||   ||||||:::|||||||||   ||||||
Db        517 GCGGCCGTGTCACTGATGGACGCCTTCCGCAGCGCCGACGAGGTGCTGCTCAACGGTGTG 576

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::||||||||||||      |||||||||||||||:::
Db        577 CAGGGCATCACCGACCTGATCACCACGCCGGGGCTGATCAACGTC 621

RESULT 15
US-10-282-122A-16538
; Sequence 16538, Application US/10282122A
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16538
;   LENGTH: 1107
;   TYPE: DNA
;   ORGANISM: Clostridium botulinum
US-10-282-122A-16538
```

```
Alignment Scores:
Pred. No.:                4.19e-46      Length:          1107
Score:                    490.00        Matches:         93
Percent Similarity:       74.86%        Conservative:    38
Best Local Similarity:    53.14%        Mismatches:      44
Query Match:              55.37%        Indels:          0
DB:                       9             Gaps:            0

US-09-770-509-2 (1-178) x US-10-282-122A-16538 (1-1107)

Qy          4 LeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeu 23
              |||:::   ||||||||||||      ||||||||   ||||||||      |||
Db        103 TTAAAAAATGTAGAGTTTATAGCTATAAATACAGATAAACAAGCTCTAATGCTTTCACAG 162

Qy         24 AlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLys 43
              |||      ||||||   :::|||      :::|||||||||||||||||||||||:::
Db        163 GCGTCTCAGAAGATACAAATAGGTGATAAATTAACAAAGGGATTAGGAGCAGGGGCTAAT 222

Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              |||||||:::||||||:::::::|||||||:::   |||   :::|||   :::   ::::::
Db        223 CCAGAAATAGGTAAGAAAGCTGCAGAAGAAAGTAAGGAAGAGATTTCTCAAGCAATAAAG 282

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
              ::::::|||:::|||||||||   ||||||||||||||||||   ||||||||||||
Db        283 GGAGCCGATATGGTATTTATAACAGCAGGTATGGGCGGAGGAACTGGTACTGGTGCAGCT 342

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||:::|||   :::|||:::   :::|||||||||||||||||||||||||:::
Db        343 CCAGTAATTGCAGAAATAGCTAAATCTATGGGAATACTTACTGTAGGTGTGGTAACTAAA 402

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||         |||      |||:::      |||
Db        403 CCTTTCCCTTTTGAAGGAAGAAAAAGATTGCTTCATGCAGAAATGGGAATAAATACTTTA 462

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
              :::   ||||||||||||:::   :::|||||:::   ||||||:::::   |||
Db        463 AAAGAGAGAGTAGACACTTTAGTTACAATCCCAAATGAAAGATTATTAAGTATAGTAGAT 522

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              |||   |||:::::::::::::::|||:::   ||||||||||||||   :::||||||
Db        523 AAAAAAACAAGCTTAATGGACTCTTTTAAACTAGCAGATGATGTTCTAAGACAAGGTGTA 582

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::::|||||||||      ||||||||||:::||||||
Db        583 CAAGGTATATCAGATTTGATTACTATACCAGGACTTGTTAATTTA 627


Search completed: June  2, 2003, 12:26:29
Job time : 5481 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 2, 2003, 08:47:53 ; Search time 1762 Seconds
(without alignments)
1636.093 Million cell updates/sec

Title: US-09-770-509-2
Perfect score: 885
Sequence: 1 ASQLEGVEFIVANTDCQALG.....LLEGVKGVTDLIVRPGLINL 178

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09770509/runat_27052003_083649_18832/app_query.fasta_1.327
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09770509_@CGN_1_1_1906_@runat_27052003_083649_18832 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		449	50.7	570	9	AI775094	AI775094 EST256194
2		438	49.5	1562	11	AY109745	AY109745 Zea mays
3		435	49.2	668	14	BQ578606	BQ578606 WHE0307_E
4		433	48.9	540	9	AJ469754	AJ469754 AJ469754
5		427	48.2	593	10	BE498211	BE498211 WHE0954_C
6		425	48.0	1537	11	AY110658	AY110658 Zea mays
7		412	46.6	627	10	AW775962	AW775962 EST335027
8		404	45.6	664	12	BG855721	BG855721 1024043F0
9		395	44.6	670	10	AW094550	AW094550 EST287730
10		379	42.8	621	13	BJ289183	BJ289183 BJ289183
11		368	41.6	615	13	BJ361253	BJ361253 BJ361253
12		363	41.0	661	13	BJ415730	BJ415730 BJ415730
13		360	40.7	663	14	BQ279894	BQ279894 1091033D0
14		358	40.5	765	14	BU012384	BU012384 QGJ1M13.y
15		352	39.8	765	14	BQ869798	BQ869798 QGD7D11.y
16		351	39.7	599	14	BQ865867	BQ865867 QGC6B17.y
17		351	39.7	731	14	BQ855906	BQ855906 QGB28A23.
18		350	39.5	654	10	AW934672	AW934672 EST353564
19		347	39.2	598	13	BI787678	BI787678 sai48d09.
20		346	39.1	426	13	BI942496	BI942496 si64d07.y
21		337	38.1	409	17	AZ302940	AZ302940 GSSBru188
22		336	38.0	541	13	BJ463421	BJ463421 BJ463421
23		332	37.5	608	13	BJ325414	BJ325414 BJ325414
24		331	37.4	605	13	BJ329349	BJ329349 BJ329349
c	25	329	37.2	544	12	BE941644	BE941644 EST421223
26		325.5	36.8	752	14	BU011950	BU011950 QGJ17J24.
27		319	36.0	492	10	AV559837	AV559837 AV559837
28		316	35.7	538	10	BE444075	BE444075 WHE1127_G
29		311	35.1	404	10	AV624700	AV624700 AV624700
30		308	34.8	458	13	BI930248	BI930248 EST550137
c	31	303.5	34.3	641	17	BH397494	BH397494 AG-ND-119
32		300	33.9	532	14	BQ279544	BQ279544 1091036G1
c	33	298.5	33.7	550	17	BH397266	BH397266 AG-ND-119
c	34	294	33.2	467	17	AZ049166	AZ049166 GSSBru056
35		294	33.2	655	14	BQ402273	BQ402273 GA__Ed004
36		292	33.0	814	10	BE643351	BE643351 Cri2_8_J1
c	37	290	32.8	448	17	BH751621	BH751621 SALK_0503
38		289	32.7	609	10	BE343445	BE343445 EST408607
c	39	288	32.5	455	17	BH751623	BH751623 SALK_0503
40		288	32.5	495	10	BE025134	BE025134 894008C08
41		287	32.4	544	13	BJ324543	BJ324543 BJ324543
42		273	30.8	468	12	BG510632	BG510632 sac71d03.
43		271	30.6	431	17	AQ863118	AQ863118 nbeb0020E
44		266.5	30.1	723	12	BG861299	BG861299 1024076G1
45		263	29.7	583	10	AV916706	AV916706 AV916706

ALIGNMENTS

RESULT 1
AI775094
LOCUS AI775094 570 bp mRNA linear EST 18-MAY-2001
DEFINITION EST256194 tomato resistant, Cornell Lycopersicon esculentum cDNA clone cLER14D3, mRNA sequence.
ACCESSION AI775094
VERSION AI775094.1 GI:5273135
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 570)

```
  AUTHORS   D' Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
            Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
            Giovannoni,J.J. and Martin,G.B.
  TITLE     Generation of ESTs from Pseudomonas resistant tomato
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Lycopersicon esculentum"
                     /cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
                     /db_xref="taxon:4081"
                     /clone="cLER14D3"
                     /clone_lib="tomato resistant, Cornell"
                     /tissue_type="leaf"
                     /dev_stage="4-week old"
                     /lab_host="SOLR"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLER - Tomato Pseudomonas Resistant EST Library.
                     Directionally cloned  cDNAs inserted into pBlueScript SK(-
                     ) at 5' end with EcoRI and 3' end  with XhoI site."
BASE COUNT      157 a    102 c    149 g    162 t
ORIGIN

Alignment Scores:
Pred. No.:              1.8e-42       Length:         570
Score:                  449.00        Matches:        90
Percent Similarity:     71.75%        Conservative:   37
Best Local Similarity:  50.85%        Mismatches:     50
Query Match:            50.73%        Indels:         0
DB:                     9             Gaps:           0

US-09-770-509-2 (1-178) x AI775094 (1-570)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::||||||:::|||         |||||||||   |||||||||   :::
Db         37 AGTGGTTTACAGGGTGTTGACTTCTATGCTATAAACACGGATGCTCAAGCACTGGTACAG 96

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||   |||   :::   :::   :::||||:::   :::|||:::||||||||||   |||
Db         97 TCTGCTGCCGAGAACCCACTTCAAATTGGAGAACTTCTGACTCGTGGGCTTGGTACTGGT 156

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
                 |||   ||||||:::::::::::||||||:::   |||       |||
Db        157 GGCAATCCTCTTTTAGGGGAACAGGCAGCAGAGGAGTCAAAGGAAGCTATTGCAAATTCT 216

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::   |||:::|||:::|||||||||   ||||||||||||||||||   :::|||
Db        217 CTGAAAGGTTCAGATATGGTTTTCATAACAGCAGGAATGGGTGGAGGTACAGGATCTGGT 276

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||||||||   :::|||:::|||   |||   ||||||||||||||||||
Db        277 GCGGCTCCTGTTGTGGCTCAAATAGCAAAAGAAGCAGGTTATTTGACTGTTGGTGTTGTT 336

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||      |||:::      |||         ::::::
Db        337 ACATATCCATTCAGCTTTGAAGGACGTAAAAGATCTGTGCAGGCTCTGGAAGCAATTGAA 396

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              :::|||   :::   ||||||||||||||||||:::||||||      ||||||   :::
Db        397 AAACTTCAGAGAAATGTTGACACACTTATAGTAATTCCCAATGATCGTCTGCTAGATATT 456

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              ||||||:::   |||   :::   :::|||||||      ||||||||||||||   :::
Db        457 GCCGATGAGCAGACACCACTTCAAGATGCTTTTCTTCTTGCGGATGATGTATTACGTCAA 516

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ||||||:::|||::::::|||:::|||      |||||||||:::|||:::
```

```
Db        517 GGTGTCCAAGGAATATCTGATATAATCACTATTCCTGGGCTTGTGAATGTG 567

RESULT 2
AY109745
LOCUS       AY109745                1562 bp    mRNA    linear   HTC 25-MAY-2002
DEFINITION  Zea mays CL421_1 mRNA sequence.
ACCESSION   AY109745
VERSION     AY109745.1  GI:21213583
KEYWORDS    HTC.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1562)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 1562)
  AUTHORS   Coe,E.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1. .1562
                     /organism="Zea mays"
                     /db_xref="MaizeDB:631962"
                     /db_xref="taxon:4577"
                     /clone="CL421_1"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      349 a    326 c    376 g    366 t    145 others
ORIGIN

Alignment Scores:
Pred. No.:              1.33e-40      Length:         1562
Score:                  438.00        Matches:        87
Percent Similarity:     71.19%        Conservative:   39
Best Local Similarity:  49.15%        Mismatches:     51
Query Match:            49.49%        Indels:         0
DB:                     11            Gaps:           0

US-09-770-509-2 (1-178) x AY109745 (1-1562)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::|||:::||||||         |||||||||   |||||||||
Db        235 AGCGGCCTCCAGGGCATCGAATTTTATGCTATAAACACCGATTCCCAAGCCCTTATTAAT 294

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||   |||   :::   :::   :::|||:::   :::|||:::||||||||||||||||
Db        295 TCACAAGCGCAATATCCTCTGCAAATTGGAGAGCAGTTGACCCGCGGCTTAGGTGCCGGT 354

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
                 |||   ||||||:::::::::::||||||:::   :::       |||
Db        355 GGAAATCCGAATTTGGGAGAGCAGGCTGCTGAGGAATCAAGAGAAACCATAGCCACTGCC 414

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::||||||:::::::::|||||||||   |||||||||       |||   :::|||
Db        415 CTGAGGGATTCAGATCTTGTCTTCATAACAGCTGGGATGGGANNNNNTACTGGATCTGGT 474

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||||||||   :::::::::|||   |||   ||||||||||||||||||
Db        475 GCTGCTCCAGTTGTTGCCCAGATATCAAAGGAAGCTGGTTATCTTACTGTTGGTGTTGTC 534

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
```

```
Db        535 ACCTATCCATTCAGTTTCGAGGGCCGTAAGCGCTCTGTACAGGCATTGGAAGCACTAGAG 594

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141

Db        595 AAGCTGGAAAAGAGTGTAGACACACTTATTGTGATTCCAAATGATAAGTTATTAGATGTT 654

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161

Db        655 GCGGATGAAAACATGCCCTTGCAAGATGCATTTCTCCTTGCAGATGATGTCCTTCGTCAG 714

Qy        162 GlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178

Db        715 GGTGTTCAAGGAATATCAGACATCATCACAATACCGGGACTTGTCAATGTT 765

RESULT 3
BQ578606
LOCUS       BQ578606                668 bp    mRNA    linear   EST 19-JUN-2002
DEFINITION  WHE0307_E03_J05ZS Wheat unstressed seedling shoot cDNA library
            Triticum aestivum cDNA clone WHE0307_E03_J05, mRNA sequence.
ACCESSION   BQ578606
VERSION     BQ578606.1  GI:21481923
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Etiolated shoot cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: SK primer.
FEATURES             Location/Qualifiers
     source          1. .668
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE0307_E03_J05"
                     /clone_lib="Wheat unstressed seedling shoot cDNA library"
                     /tissue_type="Etiolated shoot"
                     /dev_stage="Five day old seedling"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized
                     , germinated and grown aseptically in the  dark at room
                     temperature on filter paper with water, nystatin and
                     cefotaxime in covered crystallization dishes. Shoots were
                     harvested.  The tissue, total RNA, and poly(A) RNA were
                     prepared, a cDNA library  was made, and the cDNA clones
                     were in vivo excised to give pBluescript  phagemids in the
                     TJ Close lab (Choi, Close, Fenton) at the University  of
                     California, Riverside. Plasmid DNA preparations and DNA
                     sequencing  were performed in the OD Anderson lab (all
                     other authors)."
BASE COUNT      159 a    137 c    196 g    176 t
ORIGIN

Alignment Scores:
Pred. No.:              9.66e-41        Length:         668
Score:                  435.00          Matches:        81
Percent Similarity:     68.33%          Conservative:   42
Best Local Similarity:  45.00%          Mismatches:     55
Query Match:            49.15%          Indels:         2
DB:                     14              Gaps:           1

US-09-770-509-2 (1-178) x BQ578606 (1-668)

Qy          1 AlaSerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGly 20

Db          7 GCACGAGCCATGAATGGTGTCGAGTTTTGGATCGTCAACACCGATGTCCAGGCAATAAGG 66

Qy         21 ArgSer------LeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeu 38

Db         67 ATGTCCCCGGTGCATTCCCAGAACAGGCTGCAGATTGGGCAGGAGCTCACTCGGGGTTTG 126

Qy         39 GlyAlaGlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIle 58

Db        127 GGTGCGGGTGGGAACCCTGATATTGGGATGAATGCCGCCAAGGAGAGCTGTGAGTCCATA 186

Qy         59 GlnArgMetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThr 78

Db        187 GAGGAAGCTCTTCATGGTGCTGACATGGTTTTTGTCACGGCTGGAATGGGTGGGGGAACT 246

Qy         79 CysThrGlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrVal 98

Db        247 GGAACTGGAGGTGCCCCTGTAATTGCTGGAATTGCCAAGTCCATGGGTATACTAACAGTG 306

Qy         99 GlyValValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAla 118

Db        307 GGTATTGTCACAACGCCCTTTTCATTTGAGGGGAGGAGGCGGGCAGTTCAGGCTCAAGAA 366

Qy        119 GlyValLysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeu 138

Db        367 GGAATATCAGCCTTGAGAAATAGTGTGGACACTCTCATTGTCATCCCAAATGACAAGCTG 426

Qy        139 LeuAlaLeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspVal 158

Db        427 TTGTCTGCTGTTTCTCCAAATACTCCTGTCACGGAAGCATTCAACTTGGCTGATGATATT 486

Qy        159 LeuLeuGluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178

Db        487 CTTTGGCAAGGAATTCGTGGTATCTCTGATATCATTACGGTTCCTGGGTTGGTTAATGTT 546

RESULT 4
AJ469754
LOCUS       AJ469754                540 bp    mRNA    linear   EST 24-MAY-2002
DEFINITION  AJ469754 S00008 Hordeum vulgare cDNA clone S0000800253A03F1, mRNA
            sequence.
ACCESSION   AJ469754
VERSION     AJ469754.1  GI:21185710
KEYWORDS    EST.
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 540)
  AUTHORS   Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
  TITLE     Barley EST's
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Schulman AH
            Institute of Biotechnology
            University of Helsinki
            P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014,
            Finland.
FEATURES             Location/Qualifiers
     source          1. .540
                     /organism="Hordeum vulgare"
                     /db_xref="taxon:4513"
                     /clone="S0000800253A03F1"
                     /clone_lib="S00008"
                     /tissue_type="Callus"
                     /note="Callus K19"
```

```
BASE COUNT      146 a    112 c    141 g    141 t
ORIGIN

Alignment Scores:
Pred. No.:                1.25e-40       Length:         540
Score:                    433.00         Matches:        85
Percent Similarity:       72.29%         Conservative:   35
Best Local Similarity:    51.20%         Mismatches:     46
Query Match:              48.93%         Indels:         0
DB:                       9              Gaps:           0

US-09-770-509-2 (1-178) x AJ469754 (1-540)

Qy         13 AsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaProHisLysIleThrLeuGlyLys 32
Db         10 AACACAGATTCCCAGGCTCTCGTCAATTCGCAGGCACAACATCCGTTACAAATTGGAGAA 69

Qy         33 AspIleThrLysGlyLeuGlyAlaGlySerLysProGluLeuGlyLysArgSerAlaGlu 52
Db         70 CAATTGACTCGTGGGCTGGGTACTGGTGGAAATCCTAATTTGGGAGAACAGGCTGCCGAG 129

Qy         53 GlnGlnLysValAspIleGlnArgMetLeuGlnAspSerAsnMetLeuPheIleThrGly 72
Db        130 GAATCAAAGGAAGTGATAGCCAATGCCCTCCGAGATTCGGACCTTGTCTTCATAACAGCT 189

Qy         73 GlyMetGlyGlyGlyThrCysThrGlyAlaAlaProValValAlaSerValAlaArgGlu 92
Db        190 GGGATGGGAGGGGGTACTGGCTCCGGTGCTGCTCCAGTTGTTGCCCAGATAGCAAAGGAA 249

Qy         93 LeuGlyIleLeuThrValGlyValValSerThrProPheArgSerGluGlyProAsnArg 112
Db        250 GCTGGTTATCTTACTGTCGGTGTTGTCACCTACCCATTCAGCTTCGAAGGACGCAAACGC 309

Qy        113 ThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLysTyrValAspThrLeuIleVal 132
Db        310 TCTTTACAGGCACTTGAAGCATTAGAGAAGCTGGAAAGGAGTGTTGACACACTAATTGTG 369

Qy        133 ValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThrThrMetLeuGluAlaPhe 152
Db        370 ATACCAAATGATCGGTTGTTAGATATTGCTGATGAAAATATGCCCTTGCAAGACGCGTTT 429

Qy        153 ArgTyrAlaAspAspValLeuLeuGluGlyValLysGlyValThrAspLeuIleValArg 172
Db        430 CTCCTTGCAGATGATGTCCTTCGTCAGGGTGTCCAAGGAATATCAGACATAATCACGATA 489

Qy        173 ProGlyLeuIleAsnLeu 178
Db        490 CCTGGGCTTGTGAATGTT 507

RESULT 5
BE498211
LOCUS       BE498211                 593 bp    mRNA    linear   EST 04-AUG-2000
DEFINITION  WHE0954_C01_F02ZS Wheat pre-anthesis spike cDNA library Triticum
            aestivum cDNA clone WHE0954_C01_F02, mRNA sequence.
ACCESSION   BE498211
VERSION     BE498211.1  GI:9696828
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 593)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
            Seaton,C.L. and Tong,J.C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Pre-anthesis spike cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
FEATURES             Location/Qualifiers
     source          1. .593
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE0954_C01_F02"
                     /clone_lib="Wheat pre-anthesis spike cDNA library"
                     /tissue_type="Spike before anthesis"
                     /dev_stage="Adult plant"
                     /lab_host="E. coli SOLR"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
                     greenhouse.  Whole spike with awns trimmed,  white, green
                     and yellow anther were collected and total RNA, and
                     poly(A) RNA were prepared, a cDNA library was made, and
                     the cDNA  clones were in vivo excised to give pBluescript
                     phagemids in the TJ  Close lab (Choi, Close, Fenton) at
                     the University of California,  Riverside.  Plasmid DNA
                     preparations and DNA sequencing were performed  in the OD
                     Anderson lab (all other authors)."
BASE COUNT      148 a    134 c    160 g    151 t
ORIGIN

Alignment Scores:
Pred. No.:                7.12e-40       Length:         593
Score:                    427.00         Matches:        84
Percent Similarity:       71.52%         Conservative:   34
Best Local Similarity:    50.91%         Mismatches:     47
Query Match:              48.25%         Indels:         0
DB:                       10             Gaps:           0

US-09-770-509-2 (1-178) x BE498211 (1-593)

Qy          6 GlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSerLeuAlaPro 25
Db         30 GGTATCGAGTTTTATGCTATAAACACAGATTCCCAGGCTCTTGTGAATTCGCAGGCGCAA 89

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
Db         90 CATCCGCTACAAATTGGAGAGCAATTGACTCGTGGACTGGGTACTGGTGGAAATCCTAAT 149

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
Db        150 TTGGGAGAACAAGCTGCCGAGGAATCAAAGGAAGTGATAGCCAATGCCCTCCGAGATTCT 209

Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
Db        210 GATCTTGTCTTCATAACAGCTGGGATGGGAGGGGGTACTGGCTCCGGTGCTGCTCCAGTT 269

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
Db        270 GTTGCCCAGATAGCAAAGGAAGCCGGTTATCTTACTGTCGGTGTTGTCACCTACCCATTC 329

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
Db        330 AGCTTTGAAGGACGCAAGCGCTCTCTACAGGCACTCGAAGCATTGGAGAAGCTGGAAAGA 389

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
Db        390 AGTGTTGACACTCTGATTGTGATCCCAAATGATCGGTTGTTAGATATTGCTGATGAGAAT 449

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
Db        450 ATGCCCTTGCAAGACGCGTTTCTCCTTGCAGATGATGTCCTTCGACAGGGTGTCCAAGGA 509

Qy        166 ValThrAspLeuIle 170
```

```
Db        510 ATATCAGACATTATC 524

RESULT 6
AY110658
LOCUS       AY110658               1537 bp    mRNA    linear   HTC 26-MAY-2002
DEFINITION  Zea mays CL806_1 mRNA sequence.
ACCESSION   AY110658
VERSION     AY110658.1  GI:21215248
KEYWORDS    HTC.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1537)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 1537)
  AUTHORS   Coe,E.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1. .1537
                     /organism="Zea mays"
                     /db_xref="MaizeDB:632908"
                     /db_xref="taxon:4577"
                     /clone="CL806_1"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      352 a    311 c    404 g    374 t     96 others
ORIGIN

Alignment Scores:
Pred. No.:              4.33e-39        Length:         1537
Score:                  425.00          Matches:        80
Percent Similarity:     66.48%          Conservative:   39
Best Local Similarity:  44.69%          Mismatches:     58
Query Match:            48.02%          Indels:         2
DB:                     11              Gaps:           1

US-09-770-509-2 (1-178) x AY110658 (1-1537)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   :::   |||||:|||||||   :::   |||||||||   ||||||:::
Db        185 AGCTCCATGAACGGCGTTGAGTTTTGGATCGTGAACACTGATGTGCAGGCCATAAGAATG 244

Qy         22 Ser------LeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGly 39
              |||      |||      :::::::::   :::|||:::::::::|||:::|||||||||
Db        245 TCTCCTGTGCTNNNNNACAATAGACTGCAGATTGGACAGGAGTTGACTCGAGGCCTGGGC 304

Qy         40 AlaGlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGln 59
              |||         |||:::::::|||      :::|||::::::             ||||||
Db        305 GCTNNNNNAAACCCTGATATTGGGATGAATGCAGCAAAGGAGAGCAGCGAGTCCATTCAG 364

Qy         60 ArgMetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCys 79
                        :::::::|||      :::|||   ||||||||||||||||||
Db        365 GAAGCTCNNNNNGGTGCTGACATGGNNNNNGTGACGGCTGGAATGGGTGGAGGAACTGGA 424

Qy         80 ThrGlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGly 99
              ||||||   ||||||||||:::|||   :::|||:::   :::||||||||||||||||||
Db        425 ACTGGAGGTGCTCCTGTGATTGCTGGAATAGCCAAGTCCATGGGTATACTAACCGTTGGC 484

Qy        100 ValValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGly 119
              :::|||:::||||||||||      ||||||      |||         |||      |||
Db        485 ATAGTCACAACGCCTTTCTCGTTCGAGGGGAGAAGACGGGCAGTTCAAGCTCAGGAGGGA 544

Qy        120 ValLysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeu 139
              :::      |||.        ||||||||||||||||||:::||||||      ||||||
Db        545 ATAGCAGCATTGAGAAATAGTGTGGACACCCTAATCGTCATCCCAAATGATAAGTTGCTG 604

Qy        140 AlaLeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeu 159
              :::            :::|||   :::   |||||||||      ||||||||||:::|||
Db        605 TCTGCTGTTTCTCCAAATACACCTGTAACTGAAGCATTTAATCTGGCTGATGATATTCTT 664

Qy        160 LeuGluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||::::::|||:::::::|||:::|||      |||||||||:::|||:::
Db        665 CGTCAAGGCATTCGTGGCATATCTGATATAATTACGGTTCCTGGGTTGGTTAATGTT 721

RESULT 7
AW775962
LOCUS       AW775962               627 bp    mRNA    linear   EST 07-SEP-2000
DEFINITION  EST335027 DSIL Medicago truncatula cDNA clone pDSIL-3J17, mRNA
            sequence.
ACCESSION   AW775962
VERSION     AW775962.1  GI:7765775
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 627)
  AUTHORS   Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng
            ,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
            Holt,I.E. and Fraser,C.M.
  TITLE     ESTs from leaves of Medicago truncatula after inoculation with
            Colletotrichum trifolii
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Deborah A. Samac
            Department of Plant Pathology
            University of Minnesota
            495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
            Tel: 612 625 1243
            Fax: 651 649 5058
            Email: debbys@puccini.crl.umn.edu
            Minnesota sequence name:M259295e
            TIGR sequence name:MTFAK57TK
            More information is available at. .
            http://chrysie.tamu.edu/medicago
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .627
                     /organism="Medicago truncatula"
                     /cultivar="genotype A17"
                     /db_xref="taxon:3880"
                     /clone="pDSIL-3J17"
                     /clone_lib="DSIL"
                     /tissue_type="leaves infected with Colletotrichum
                     trifolii"
                     /dev_stage="cotyledons and primary leaves harvested 5 and
                     8 days after inoculation with Colletotrichum trifolii"
                     /lab_host="E. coli strain XLOLR"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA from
                     cotyledons and primary leaves harvested 5 and 8 days after
                     inoculation with Colletotrichum trifolii. The cDNA was
                     directionally ligated into the Uni-ZAP XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-ZAP phage using Ex-Assist
                     helper phage and propagated in XLOLR cells. Note: EST may
                     be of fungal origin."
BASE COUNT      166 a    112 c    161 g    187 t      1 others
ORIGIN
```

```
Alignment Scores:
Pred. No.:                4.37e-38        Length:         627
Score:                    412.00          Matches:        84
Percent Similarity:       70.41%          Conservative:   35
Best Local Similarity:    49.70%          Mismatches:     50
Query Match:              46.55%          Indels:         0
DB:                       10              Gaps:           0

US-09-770-509-2 (1-178) x AW775962 (1-627)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::|||||||::|||         |||||||||  |||||||||
Db        118 AGTGGTTTGCAGGGTGTAGACTTCTATGCAATAAATACTGATGCTCAAGCACTACTACAT 177

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||   |||   :::   |||   :::|||:::   :::|||:::||||||||||   |||
Db        178 TCAGCTGCTGAGAATCCTATTAAAATCGGAGAGCTTCTGACTCGTGGATTAGGTACCGGT 237

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
                    |||   ||||||:::::::::::||||||:::   |||       |||
Db        238 GGGAATCCACTTTTGGGCGAACAAGCTGCCGAGGAATCAAAAGAAGCTATTGCTGATGCC 297

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::   |||:::::::::|||||||||   ||||||||||||||||||   :::|||
Db        298 CTTAAAGGATCAGATTTGGTGTTTATAACAGCTGGGATGGGTGGGGGTACAGGGTCTGGT 357

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||||||||   :::::::::|||   |||   ||||||||||||||||||
Db        358 GCTGCACCAGTTGTGGCCCAAATATCAAAAGAGGCAGGTTACTTGACTGTAGGTGTTGTT 417

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||       ||||||       |||:::       |||          ::::::
Db        418 ACATATCCTTTCAGTTTTGAAGGACGTAAAAGATCCTTGCAGGCACTTGAAGCCATTGAA 477

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              :::|||   :::   ||||||||||||||||||:::||||||     ||||||   :::
Db        478 AAGCTTCAGAGAAATGTTGATACACTTATTGTAATTCCAAATGATCGTCTGCTTGACATA 537

Qy        142 AlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGlu 161
              ||||||:::       :::   :::||||||||||   ||||||||||||||   :::
Db        538 GCTGACGAGCAGATGCCCCTCCAGGATGCTTTTCGTCTTGCAGATGATGTTTTACGCCAA 597

Qy        162 GlyValLysGlyValThrAspLeuIle 170
              ||||||   |||:::::::::::|||
Db        598 GGAGTTCANGGAATTTCAAACATTATA 624

RESULT 8
BG855721
LOCUS       BG855721                 664 bp    mRNA    linear   EST 29-MAY-2001
DEFINITION  1024043F01.y2 C. reinhardtii CC-1690, normalized, Lambda Zap II
            Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BG855721
VERSION     BG855721.1  GI:14236905
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
            McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants; project phase 2
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1. .664
                     /organism="Chlamydomonas reinhardtii"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
                     II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; This library, constructed by John Davies and Jeffrey
                     McDermott, combines cDNAs from CC-1690 cells grown to
                     mid-log phase in TAP (acetate-containing) medium in the
                     light, TAP medium in the dark, HS (minimal) medium in
                     ambient levels of CO2 and HS medium bubbled with 5% CO2.
                     PolyA mRNA was purified from each sample, pooled and cDNA
                     synthesized. The cDNA was directionally cloned into lambda
                     ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
                     pBluescript II SK- plasmids were excised from the lambda
                     ZAP clones by superinfection with ExAssist (Stratagene)
                     phage. The library was normalized using method 4 described
                     in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT      123 a    209 c    204 g    128 t
ORIGIN

Alignment Scores:
Pred. No.:                4.08e-37        Length:         664
Score:                    404.00          Matches:        83
Percent Similarity:       68.75%          Conservative:   27
Best Local Similarity:    51.88%          Mismatches:     48
Query Match:              45.65%          Indels:         2
DB:                       12              Gaps:           1

US-09-770-509-2 (1-178) x BG855721 (1-664)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   ::::::|||||||||||   :::|||||||||||   |||||||||
Db        185 AGCGACGTGCAGGGCGTGGAGTTCTGGATTGCCAACACCGACGCTCAGGCCTTGGCAACG 244

Qy         22 Ser------LeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGly 39
              |||                  |||:::   :::|||     :::|||:::|||||||||
Db        245 TCTCCCGTTAACGGCAAGTGCAAGGTGCAAATTGGAGGCAAGCTTACCCGCGGTCTCGGC 304

Qy         40 AlaGlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGln 59
              ||||||       ||||||:::|||   ::::::||||||:::   :::     |||
Db        305 GCCGGAGGCAACCCTGAGATCGGCGCTAAAGCTGCTGAAGAGAGCCGGGACTCCATCGCC 364

Qy         60 ArgMetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCys 79
                    ||||||||||:::::|||:::|||:::|||   ||||||||||||||||||
Db        365 GCGGCACTGCAAGATACTGACATGGTATTCGTGACGGCCGGAATGGGCGGCGGCACGGGC 424

Qy         80 ThrGlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGly 99
              :::|||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        425 AGTGGCGCCGCGCCCGTCGTGGCGCAGGTGGCGCGTGAATTGGGCATCCTAACAGTTGGC 484

Qy        100 ValValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGly 119
              :::|||:::|||||||||       ||||||       |||   :::   |||   :::
Db        485 ATCGTCACCACCCCCTTCACCTTCGAGGGCCGCCAGCGCGCGCAGCAGGCTCGCTCTGCT 544

Qy        120 ValLysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeu 139
              :::       |||           |||||||||||||||||:::||||||       ||||||
Db        545 TTAGCCAACTTGCGTGCAGCGGTTGACACGCTCATTGTCATCCCGAACGACCGGCTGCTG 604

Qy        140 AlaLeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeu 159
              :::       |||   :::       :::   |||       :::   ||||||||||||||
Db        605 TCGGCCATGGACTCCAACGTGCCTATCAAGGAGCGCCTTCAAATTGCCGATGACGTACTG 664

RESULT 9
AW094550
LOCUS       AW094550                 670 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST287730 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
            clone cLET29K5, mRNA sequence.
ACCESSION   AW094550
VERSION     AW094550.1  GI:6060145
KEYWORDS    EST.
SOURCE      tomato.
```

```
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 670)
  AUTHORS   D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
            Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
            Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and Giovannoni
            ,J.
  TITLE     Generation of ESTs from tomato leaf tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .670
                     /organism="Lycopersicon esculentum"
                     /cultivar="Rio Grande PtoR"
                     /db_xref="taxon:4081"
                     /clone="cLET29K5"
                     /clone_lib="tomato mixed elicitor, BTI"
                     /tissue_type="leaf"
                     /dev_stage="4-6 week old plants"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLET - Inoculated with a variety of disease response
                     elicitors. Plants  exposed to 2,6 dichloroisonicotinic
                     acid, BTH, jasmonic acid, ethylene,  fenthion, EIX,
                     okadaic acid, or systemin prior to tissue harvest.  EcoRI
                     site was destroyed during cloning."
BASE COUNT      187 a    114 c    181 g    188 t
ORIGIN

Alignment Scores:
Pred. No.:                4.67e-36       Length:         670
Score:                    395.00         Matches:        71
Percent Similarity:       71.24%         Conservative:   38
Best Local Similarity:    46.41%         Mismatches:     44
Query Match:              44.63%         Indels:         0
DB:                       10             Gaps:           0

US-09-770-509-2 (1-178) x AW094550 (1-670)

Qy         26 HisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySerLysProGlu 45
              :::::::::   :::|||:::   :::|||:::|||||||||||||||     |||::: 
Db          5 AATCGATTGCCAATAGGCCAATAGCTCACGAGAGGACTAGGTGCAGGTGGTAATCCAGAT 64

Qy         46 LeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGlnAspSer 65
              ::::||      :::|||::::::   |||      |||:::      ::::::   ::: 
Db         65 ATAGGGATGAATGCTGCCAAAGAAAGCAAGGAGGCTATCGAAGAAGCAGTTCGCGGTGCA 124

Qy         66 AsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAlaProVal 85
              :::|||:::|||:::|||   ||||||||||||||||||   ||||||   ||||||::: 
Db        125 GATATGGTTTTTGTGACTGCTGGAATGGGCGGAGGAACAGGGACTGGTGGGGCTCCTATA 184

Qy         86 ValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThrProPhe 105
              :::|||   :::|||:::   :::|||||||||||||||||||:::|||:::|||||||| 
Db        185 ATTGCAGGAATTGCCAAATCAATGGGTATCTTAACTGTTGGTATTGTCACAACCCCCTTT 244

Qy        106 ArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeuAlaLys 125
              ||||||      |||          |||      |||:::      |||   ::: 
Db        245 TCTTTTGAGGGACGAAGAAGAGCAGTTCAAGCCCAAGAAGGAATTGCAGCTTTGAGAGAA 304

Qy        126 TyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAspLysSer 145
              ||||||||||||||||||:::||||||      ||||||                  ||| 
Db        305 AATGTTGATACACTAATTGTCATTCCTAATGACAAGTTACTGACTGCTGTTTCCTTATCA 364

Qy        146 ThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyValLysGly 165
              |||   :::   |||||||||      ||||||||||:::|||   :::||||||:::||| 
```

```
Db        365 ACCCCAGTAACTGAAGCTTTTAACCTGGCTGATGATATTCTTCGGCAAGGGGTTCGTGGT 424

Qy        166 ValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              ::::::|||:::|||      ||||||||||:::|||::: 
Db        425 ATTTCTGATATAATTACGATTCCTGGACTGGTAAATGTT 463

RESULT 10
BJ289183
LOCUS       BJ289183                621 bp    mRNA    linear   EST 09-APR-2002
DEFINITION  BJ289183 Y. Ogihara unpublished cDNA library, Wh_SL Triticum
            aestivum cDNA clone whsl3n10 5', mRNA sequence.
ACCESSION   BJ289183
VERSION     BJ289183.1  GI:20107251
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   Ogihara,Y. and Murai,K.
  TITLE     Expressed genes in Triticum aestivum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="whsl3n10"
                     /clone_lib="Y. Ogihara unpublished cDNA library, Wh_SL"
                     /tissue_type="seed DPA30"
                     /dev_stage="Feekes' scale 11.3"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown under
                     hydroponic conditions at UC Davis, salt stressed  for 12
                     hours, and for 7 days, then dissected and frozen (Akhunov
                     in J  Dvorak Lab). Total RNA was prepared from sheath
                     tissue, equal quantities  of RNA were pooled from the two
                     samples, polyA was purified from the  pooled RNA, a cDNA
                     library was made, and the cDNA clones were in vivo
                     excised to give pBluescript phagemids in the TJ Close lab
                     at the  University of California, Riverside (Akhunov, Chin
                     , Choi, Close, Fenton,  Kianian, Otto, Simons, Zhang).
                     Plasmid DNA preparations and DNA  sequencing were
                     performed in the OD Anderson lab (all other authors)."
BASE COUNT      150 a    120 c    187 g    163 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:                3.15e-34       Length:         621
Score:                    379.00         Matches:        69
Percent Similarity:       69.44%         Conservative:   31
Best Local Similarity:    47.92%         Mismatches:     44
Query Match:              42.82%         Indels:         0
DB:                       13             Gaps:           0

US-09-770-509-2 (1-178) x BJ289183 (1-621)

Qy         35 ThrLysGlyLeuGlyAlaGlySerLysProGluLeuGlyLysArgSerAlaGluGlnGln 54
              |||:::|||||||||||||||      |||:::::::|||      :::|||:::::: 
Db         20 ACTCGGGGTCTGGGTGCGGGTGGGAACCCTGATATTGGGATGAATGCTGCCAAGGAGAGC 79

Qy         55 LysValAspIleGlnArgMetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMet 74
                    |||:::      |||      ::::::|||:::|||:::|||   ||||||
Db         80 TGTGAGTCCATAGAGGAAGCACTTCATGGTGCTGACATGGTTTTTGTCACGGCAGGAATG 139
```

```
Qy          75 GlyGlyGlyThrCysThrGlyAlaAlaProValValAlaSerValAlaArgGluLeuGly 94
               ||||||||||||  ||||||  |||||||||:::|||  :::|||:::  :::|||
Db         140 GGTGGGGGAACTGGAACTGGAGGTGCCCCTGTAATTGCTGGAATTGCCAAGTCCATGGGT 199

Qy          95 IleLeuThrValGlyValValSerThrProPheArgSerGluGlyProAsnArgThrArg 114
               ||||||||||||||:::|||:::|||||||||      ||||||      |||
Db         200 ATACTGACAGTGGGTATTGTCACAACGCCCTTTTCATTTGAGGGGAGGAGGCGGGCAGTT 259

Qy         115 LeuAlaAsnAlaGlyValLysGluLeuAlaLysTyrValAspThrLeuIleValValPro 134
                |||      |||:::      |||      ||||||||||||||||||:::|||
Db         260 CAGGCTCAAGAAGGAATATCAGCCTTGAGAAATAGTGTGGACACTCTCATTGTCATCCCA 319

Qy         135 AsnGlnAsnLeuLeuAlaLeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyr 154
               |||      ||||||:::           :::|||  :::  |||||||||
Db         320 AATGACAAGCTGTTGTCTGCTGTTTCTCCAAACACTCCTGTCACGGAAGCATTCAACTTG 379

Qy         155 AlaAspAspValLeuLeuGluGlyValLysGlyValThrAspLeuIleValArgProGly 174
               |||||||||:::|||  :::|||:::::|||:::::|||:::|||      ||||||
Db         380 GCTGATGATATTCTTTGGCAAGGAATTCGCGGTATCTCTGATATCATTACGGTTCCTGGG 439

Qy         175 LeuIleAsnLeu 178
               |||:::|||:::
Db         440 CTGGTTAATGTT 451

RESULT 11
BJ361253
LOCUS       BJ361253                615 bp    mRNA    linear   EST 07-MAR-2002
DEFINITION  BJ361253 Dictyostelium discoideum cDNA library, CF Dictyostelium
            discoideum cDNA clone ddc16n12 5', mRNA sequence.
ACCESSION   BJ361253
VERSION     BJ361253.1  GI:19260848
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 615)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the culmination
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .615
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddc16n12"
                     /clone_lib="Dictyostelium discoideum cDNA library, CF"
                     /sex="mat A"
                     /dev_stage="Culmination stage"
BASE COUNT      242 a     77 c    124 g    171 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              6.03e-33        Length:         615
Score:                  368.00          Matches:        70
Percent Similarity:     72.26%          Conservative:   29
Best Local Similarity:  51.09%          Mismatches:     38
Query Match:            41.58%          Indels:         0
DB:                     13              Gaps:           0

US-09-770-509-2 (1-178) x BJ361253 (1-615)

Qy           3 GlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSer 22
               :::|||   |||:::::::|||:::|||||||||||||||   |||||||||      |||
Db         196 GAATTATATGGAATTGATTTTGTAGTTGCCAATACTGATGCACAAGCATTGGCAATATCA 255

Qy          23 LeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySer 42
                 :::          :::   |||||||||   :::|||:::|||||||||||||||:::
Db         256 TGTAGTAGAAAGATGGTACAATTAGGAAAGACATTAACAAGAGGATTAGGAGCAGGAGCA 315

Qy          43 LysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeu 62
               ||||||:::|||||||::::::   |||:::         ::::::         :::
Db         316 GTACCAGAAGTTGGAAAGAAAGCAACTGAAGAATCAATTGAAGAATTAATGAATCAAATT 375

Qy          63 GlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAla 82
                |||:::   |||||||||:::|||   ||||||||||||||||   ||||||
Db         376 GGTGATACACAAATGTTATTTGTCACAGCTGGTATGGGTGGTGGTACAGGTACAGGTGGA 435

Qy          83 AlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSer 102
               |||   |||:::||||||   |||:::      |||||||||||||||||:::|||:::
Db         436 GCAGCAGTTATTGCATCAGCAGCAAAAGCCAAAGGTATTTTAACTGTTGGTATTGTAACC 495

Qy         103 ThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGlu 122
               ||||||      ||||||   :::|||   :::||||||      |||:::   |||
Db         496 AAACCATTTCATTTCGAAGGTAAACATAGAATGAAATTGGCAGAACAAGGTTTGATAGAG 555

Qy         123 LeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeu 139
               |||   |||   ||||||:::|||||||||:::|||||||:::   |||:::
Db         556 TTGGAGAAATCAGTCGATAGTTTAATCGTTATTCCAAATGAGAAATTAATG 606

RESULT 12
BJ415730
LOCUS       BJ415730                661 bp    mRNA    linear   EST 10-MAR-2002
DEFINITION  BJ415730 Dictyostelium discoideum cDNA library, VF Dictyostelium
            discoideum cDNA clone ddv23d18 5', mRNA sequence.
ACCESSION   BJ415730
VERSION     BJ415730.1  GI:19328123
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 661)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the vegetative
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .661
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="ddv23d18"
                     /clone_lib="Dictyostelium discoideum cDNA library, VF"
                     /sex="mat A"
                     /dev_stage="Growth phase"
BASE COUNT      257 a     87 c    134 g    182 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.55e-32        Length:         661
Score:                  363.00          Matches:        69
Percent Similarity:     71.53%          Conservative:   29
Best Local Similarity:  50.36%          Mismatches:     39
Query Match:            41.02%          Indels:         0
DB:                     13              Gaps:           0

US-09-770-509-2 (1-178) x BJ415730 (1-661)

Qy           3 GlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArgSer 22
               :::|||   |||:::::::|||:::|||||||||||||||   |||||||||      |||
Db         241 GAATTATATGGAATTGATTTTGTAGTTGCCAATACTGATGCACAAGCATTGGCAATATCA 300
```

```
Qy         23 LeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGlySer 42
              :::          :::   |||||||||   :::|||:::||||||||||||||||:::
Db        301 TGTAGTAGAAAGATGGTACAATTAGGAAAGACATTAACAAGAGGATTAGGAGCAGGAGCA 360

Qy         43 LysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeu 62
              ||||||:::|||||:::::   |||:::         ::::::         :::
Db        361 GTACCAGAAGTTGGAAAGAAAGCAACTGAAGAATCAATTGAAGAATTAATGAATCAAATT 420

Qy         63 GlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAla 82
              |||:::  ||||||||||:::|||  ||||||||||||||||  ||||||
Db        421 GGTGATACACAAATGTTATTTGTCACAGCTGGTATGGGTGGTGGTACAGGTACAGGTGGA 480

Qy         83 AlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSer 102
              |||   |||:::||||||   |||:::      ||||||||||||||||||:::|||:::
Db        481 GCAGCAGTTATTGCATCAGCAGCAAAAGCCAAAGGTATTTTAACTGTTGGTATTGTAACC 540

Qy        103 ThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGlu 122
              ||||||      ||||||   :::|||   ::::||||||      |||:::   |||
Db        541 AAACCATTTCATTTCGAAGGTAAACATAGAATGAAATTGGCAGAACAAGGTTTGATAGAG 600

Qy        123 LeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeu 139
              |||   |||      |||:::|||||||||:::||||||:::   |||:::
Db        601 TTGGAGAAATCAGNCGATAGTTTAATCGTTATTCCAAATGAGAAATTAATG 651

RESULT 13
BQ279894
LOCUS       BQ279894                663 bp    mRNA    linear   EST 08-MAY-2002
DEFINITION  1091033D05.y1 1091 - Immature ear with common ESTs screened by
            Schmidt lab Zea mays cDNA, mRNA sequence.
ACCESSION   BQ279894
VERSION     BQ279894.1  GI:20508091
KEYWORDS    EST.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 663)
  AUTHORS   Walbot,V.
  TITLE     Maize ESTs from various cDNA libraries sequenced at Stanford
            University
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 1091033  row: D  column: 05.
FEATURES             Location/Qualifiers
     source          1. .663
                     /organism="Zea mays"
                     /cultivar="OH43"
                     /db_xref="taxon:4577"
                     /clone_lib="1091 - Immature ear with common ESTs screened
                     by Schmidt lab"
                     /tissue_type="Inflorescence meristem - floral organ
                     primordia"
                     /dev_stage="0.5 cm to 2 cm"
                     /lab_host="Stratagene XLOLR"
                     /note="Organ: Immature ear; Vector: pAD-GAL4; Site_1:
                     EcoRI; Site_2: XhoI; RNA from library 606 was filtered for
                     common ESTs found in 606."
BASE COUNT      147 a    172 c    196 g    148 t
ORIGIN

Alignment Scores:
Pred. No.:              5.75e-32        Length:         663
Score:                  360.00          Matches:        72
Percent Similarity:     70.83%          Conservative:   30
Best Local Similarity:  50.00%          Mismatches:     42
Query Match:            40.68%          Indels:         0

DB:                     14              Gaps:           0

US-09-770-509-2 (1-178) x BQ279894 (1-663)

Qy          2 SerGlnLeuGluGlyValGluPheIleValAlaAsnThrAspCysGlnAlaLeuGlyArg 21
              |||   |||:::|||:::||||||         |||||||||   |||||||||
Db        217 AGCGGCCTCCAGGGCATCGAATTTTATGCTATAAACACCGATTCCCAAGCCCTTATTAAT 276

Qy         22 SerLeuAlaProHisLysIleThrLeuGlyLysAspIleThrLysGlyLeuGlyAlaGly 41
              |||   |||   :::   :::   :::|||:::   :::|||:::||||||||||   |||
Db        277 TCACAAGCGCAATATCCTCTGCAAATTGGAGAGCAGTTGACCCGCGGCTTAGGTACTGGT 336

Qy         42 SerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMet 61
                  |||   ||||||:::::::::|||||||:::   :::       |||
Db        337 GGAAATCCGAATTTGGGAGAGCAGGCTGCTGAGGAATCAAGAGAAACCATAGCCACTGCC 396

Qy         62 LeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGly 81
              |||:::||||||:::::::::|||||||||   ||||||||||||||||||   :::|||
Db        397 CTGAGGGATTCAGATCTTGTCTTCATAACAGCTGGGATGGGAGGGGGTACTGGATCTGGT 456

Qy         82 AlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValVal 101
              ||||||||||||||||||   ::::::::::|||   |||   ||||||||||||||||||
Db        457 GCTGCTCCAGTTGTTGCCCAGATATCAAAGGAAGCTGGTTATCTTACTGTTGGTGTTGTC 516

Qy        102 SerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLys 121
              :::   ||||||      ||||||      |||:::      |||          ::::::
Db        517 ACCTATCCATTCAGTTTCGAGGGCCGTAAGCGCTCTGTACAGGCATTGGAAGCACTAGAG 576

Qy        122 GluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeu 141
              :::|||   |||   ||||||||||||||||||:::||||||      ||||||   :::
Db        577 AAGCTGGAAAAGAGTGTAGACACACTTATTGTGATTCCAAATGATAAGTTATTAGATGTT 636

Qy        142 AlaAspLysSer 145
              ||||||::::::
Db        637 GCGGATGAAAAC 648

RESULT 14
BU012384
LOCUS       BU012384                765 bp    mRNA    linear   EST 22-AUG-2002
DEFINITION  QGJ1M13.yg.ab1 QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
            QGJ1M13, mRNA sequence.
ACCESSION   BU012384
VERSION     BU012384.1  GI:22446779
KEYWORDS    EST.
SOURCE      Lactuca sativa.
  ORGANISM  Lactuca sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QG_CA_Contig7869, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QGJ1  row: M  column: 13.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Lactuca sativa"
                     /cultivar="L.serriola"
                     /db_xref="taxon:4236"
```

/clone="QGJ1M13"
/clone_lib="QG_EFGHJ lettuce serriola"
/lab_host="E.coli"
/note="Vector: pBRcDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/ TAG_SEQ=Not found"

BASE COUNT 208 a 153 c 208 g 196 t
ORIGIN

Alignment Scores:

Pred. No.:	1.19e-31	Length:	765
Score:	358.00	Matches:	71
Percent Similarity:	72.46%	Conservative:	29
Best Local Similarity:	51.45%	Mismatches:	38
Query Match:	40.45%	Indels:	0
DB:	14	Gaps:	0

US-09-770-509-2 (1-178) x BU012384 (1-765)

```
Qy         41 GlySerLysProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArg 60
              |||      |||   |||||::::::::::|||||:::   |||       |||
Db         21 GGTGGGAATCCGATGTTGGGGGAACAGGCTGCAGAGGAATCAAAGGAAGCAATTGGAAAT 80

Qy         61 MetLeuGlnAspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThr 80
                 |||:::   |||::::::::::|||||||||   ||||||||||||||||||   :::
Db         81 GCTCTCGAGGGATCAGATCTTGTGTTCATAACAGCAGGAATGGGAGGTGGGACAGGATCT 140

Qy         81 GlyAlaAlaProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyVal 100
              ||||||||||||||||||||   :::|||:::|||   |||   ||||||||||||.|||
Db        141 GGAGCTGCTCCAGTGGTGGCCCAGATAGCAAAAGAAGCAGGATATCTCACCGTTGGTGTT 200

Qy        101 ValSerThrProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyVal 120
              |||:::   ||||||      ||||||      |||:::      |||       :::
Db        201 GTAACCTACCCCTTCAGCTTTGAAGGACGCAAAAGATCAGTCCAGGCTTTGGAAGCTATT 260

Qy        121 LysGluLeuAlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAla 140
              :::::::|||   :::   ||||||||||||||||||:::|||||||      ||||||
Db        261 GAGAAGCTGCAAAGGAATGTGGACACACTTATAGTCATCCCAAATGACCGCTTACTGGAT 320

Qy        141 LeuAlaAspLysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeu 160
              :::|||||:::   |||   :::   :::||||||      ||||||||||||||||
Db        321 ATTGCTGATGAGCAGACACCCCTTCAGGATGCTTTTCTTCTTGCTGATGATGTACTCCGT 380

Qy        161 GluGlyValLysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::||||||:::|||:::::::|||:::|||      |||||||||:::|||:::
Db        381 CAAGGCGTTCAAGGCATTTCAGATATAATCACGATACCTGGGCTTGTGAACGTG 434
```

RESULT 15
BQ869798

LOCUS	BQ869798 765 bp mRNA linear EST 15-AUG-2002
DEFINITION	QGD7D11.yg.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone QGD7D11, mRNA sequence.
ACCESSION	BQ869798
VERSION	BQ869798.1 GI:22255585
KEYWORDS	EST.
SOURCE	Lactuca sativa.
ORGANISM	Lactuca sativa Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactuca.
REFERENCE	1 (bases 1 to 765)
AUTHORS	Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE	Lettuce and Sunflower ESTs from the Compositae Genome Project http://compgenomics.ucdavis.edu/
JOURNAL	Unpublished (2002)
COMMENT	Contact: Alexander Kozik [R.W.Michelmore] Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Asmundson Hall, UCD, Davis, CA 95616, USA Tel: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu] belongs to contig QG_CA_Contig7869, see http://cgpdb.ucdavis.edu/ for details. Plate: QGD7 row: D column: 11.

FEATURES Location/Qualifiers
source 1. .765
/organism="Lactuca sativa"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QGD7D11"
/clone_lib="QG_ABCDI lettuce salinas"
/lab_host="E.coli"
/note="Vector: pBRcDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/ TAG_LIB=QG_ABCDI lettuce salinas
TAG_TISSUE=chemical induction
TAG_SEQ=TGTAGCCGGG"

BASE COUNT 213 a 154 c 202 g 195 t 1 others
ORIGIN

Alignment Scores:

Pred. No.:	6.01e-31	Length:	765
Score:	352.00	Matches:	70
Percent Similarity:	73.33%	Conservative:	29
Best Local Similarity:	51.85%	Mismatches:	36
Query Match:	39.77%	Indels:	0
DB:	14	Gaps:	0

US-09-770-509-2 (1-178) x BQ869798 (1-765).

```
Qy         44 ProGluLeuGlyLysArgSerAlaGluGlnGlnLysValAspIleGlnArgMetLeuGln 63
              |||   ||||||::::::::::|||||:::   |||       |||          |||:::
Db          8 CCGATGTTGGGGGAACAGGCTGCAGAGGAATCAAAGGAAGCAATCGGAAATGCTCTCGAG 67

Qy         64 AspSerAsnMetLeuPheIleThrGlyGlyMetGlyGlyGlyThrCysThrGlyAlaAla 83
                 |||::::::::::|||||||||   ||||||||||||||||||   :::|||||||||
Db         68 GGATCAGATCTTGTGTTCATAACAGCAGGAATGGGAGGTGGGACAGGATCTGGAGCTGCT 127

Qy         84 ProValValAlaSerValAlaArgGluLeuGlyIleLeuThrValGlyValValSerThr 103
              ||||||||||||   :::|||:::|||   |||   ||||||||||||||||||:::
Db        128 CCAGTGGTGGCCCAGATAGCAAAAGAAGCAGGATATCTCACCGTTGGTGTTGTAACCTAC 187

Qy        104 ProPheArgSerGluGlyProAsnArgThrArgLeuAlaAsnAlaGlyValLysGluLeu 123
              ||||||      ||||||      |||:::      |||         ::::::::::|||
Db        188 CCCTTCAGCTTTGAAGGACGCAAAAGATCAGTCCAGGCTTTGGAAGCTATTGAGAAGCTG 247

Qy        124 AlaLysTyrValAspThrLeuIleValValProAsnGlnAsnLeuLeuAlaLeuAlaAsp 143
                 :::   ||||||||||||||||||:::||||||      ||||||   :::||||||
Db        248 CAAAGGAATGTGGACACACTTATAGTAATCCCAAATGACCGCTTACTGGATATTGCTGAT 307

Qy        144 LysSerThrThrMetLeuGluAlaPheArgTyrAlaAspAspValLeuLeuGluGlyVal 163
              :::   |||   :::   :::||||||      ||||||||||||||   :::||||||
Db        308 GAGCAGACACCCCTTCAGGATGCTTTTCTTCTTGCTGATGATGTACTCCGTCAAGGCGTT 367

Qy        164 LysGlyValThrAspLeuIleValArgProGlyLeuIleAsnLeu 178
              :::|||:::::::|||:::|||      |||||||||:::|||:::
Db        368 CAAGGCATTTCAGATATAATCACGATACCTGGGCTTGTGAACGTG 412
```

Search completed: June 2, 2003, 10:07:12
Job time : 1787 secs